Advances in Mathematics 421 (2023) 108994

Contents lists available at ScienceDirect

Advances in Mathematics

journal homepage: www.elsevier.com/locate/aim

The shrinking target problem for matrix transformations of tori: Revisiting the standard problem

Bing Li [a,1], Lingmin Liao [b], Sanju Velani [c,*], Evgeniy Zorin [c]

[a] *Department of Mathematics, South China University of Technology, Wushan Road 381, Tianhe District, Guangzhou, China*
[b] *School of Mathematics and Statistics, Wuhan University, Bayi Road 299, Wuchang District, Wuhan, China*
[c] *Department of Mathematics, University of York, Heslington, York, YO10 5DD, England, United Kingdom of Great Britain and Northern Ireland*

A R T I C L E   I N F O

Article history:
Received 19 January 2023
Accepted 13 March 2023
Available online 6 April 2023
Communicated by Lars Olsen

MSC:
primary 11J83, 28D05
secondary 11K60, 28A80, 37A44, 60F20

Keywords:
Quantitative zero-one laws
Dynamical systems
Metric Diophantine approximation
Hausdorff dimension

A B S T R A C T

Let T be a $d \times d$ matrix with real coefficients. Then T determines a self-map of the d-dimensional torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of subsets of $\mathbb{T}^d$ and let $W(T, \{E_n\})$ be the set of points $\mathbf{x} \in \mathbb{T}^d$ such that $T^n(\mathbf{x}) \in E_n$ for infinitely many $n \in \mathbb{N}$. For a large class of subsets (namely, those satisfying the so called bounded property (**B**) which includes balls, rectangles, and hyperboloids) we show that the d-dimensional Lebesgue measure of the shrinking target set $W(T, \{E_n\})$ is zero (resp. one) if a natural volume sum converges (resp. diverges). In fact, we prove a quantitative form of this zero-one criteria that describes the asymptotic behaviour of the counting function $R(x, N) := \#\{1 \le n \le N : T^n(x) \in E_n\}$. The counting result makes use of a general quantitative statement that holds for a large class measure-preserving dynamical systems (namely, those satisfying the so called summable-mixing property). We next turn our attention to the Hausdorff dimension of $W(T, \{E_n\})$. In the case the subsets E_n are balls, rectangles or hyperboloids we obtain precise formulae for the dimension.

* Corresponding author.
E-mail addresses: scbingli@scut.edu.cn (B. Li), lmliao@whu.edu.cn (L. Liao), sanju.velani@york.ac.uk (S. Velani), evgeniy.zorin@york.ac.uk (E. Zorin).
[1] BL was supported partially by NSFC 12271176.

https://doi.org/10.1016/j.aim.2023.108994

These shapes correspond, respectively, to the simultaneous, weighted and multiplicative theories of classical Diophantine approximation. The dimension results for balls generalises those obtained in [25] for integer matrices to real matrices. In the final section, we discuss various problems that stem from the results proved in the paper.

"You've always had the power my dear,
you just had to learn it yourself."

1. Introduction

Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving dynamical system. Recall, that by definition μ is a probability measure. Now let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of subsets in $\mathcal{B}$ and let

$$\begin{aligned} W\big(T, \{E_n\}\big) &:= \limsup_{n\to\infty} T^{-n}(E_n) \\ &= \{x \in X : T^n(x) \in E_n \ \text{ for infinitely many } n \in \mathbb{N}\}. \end{aligned}$$

For obvious reasons the sets E_n can be thought of as targets that the orbit under T of points in X have to hit. The interesting situation is usually, when working within a metric space, the diameters of E_n tend to zero as n increases. It is thus natural to refer to $W\big(T, \{E_n\}\big)$ as the corresponding shrinking target set associated with the given dynamical system and target sets. Since T is measure-preserving $\mu(T^{-n}(E_n)) = \mu(E_n)$, and a straightforward consequence of the (convergent) Borel-Cantelli Lemma is that

$$\mu\big(W(T, \{E_n\})\big) = 0 \qquad \text{if} \qquad \sum_{n=1}^{\infty} \mu(E_n) < \infty. \tag{1}$$

Now two natural questions arise. Both fall under the umbrella of the "shrinking target problem" formulated in [24].

- **(P1)** What is the μ-measure of $W(T, \{E_n\})$ if the measure sum in (1) diverges?
- **(P2)** What is the Hausdorff dimension of $W(T, \{E_n\})$ if the measure sum converges and so $\mu\big(W(T, \{E_n\})\big) = 0$?

To be precise, the target sets E_n in the original formulation in [24] are restricted to balls B_n. The more general setup naturally incorporates a larger class of problems. For example, within the context of simultaneous Diophantine approximation, it enables us to address problems associated with the weighted (the target sets are rectangular) and multiplicative (the target sets are hyperbola) theories – see Remark 4 in §1.2 below.

In this paper we revisit the shrinking target problem investigated in [25] in which T is a matrix transformation of the d-dimensional torus $X = \mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. There are several reasons for doing this. Firstly, for integer matrix transformations a solution to (P1) was announced in [25]; namely, under some regularity condition on the rate at which the diameters of the balls B_n tend to zero, we have that

$$m_d(W(T, \{B_n\})) = 1 \qquad \text{if} \qquad \sum_{n=1}^{\infty} m_d(B_n) = \infty \tag{2}$$

where m_d is d-dimensional Lebesgue measure. However, the intended paper establishing this divergent analogue of (1) was never completed[2] and to the best of our knowledge such a result has not to date appeared in print elsewhere. In this paper not only do we rectify the situation but we consider the set up in which T is a real (rather than just integer) matrix transformation and the 'target' sets are general sets rather than just balls. Furthermore, our results are significantly stronger than statements such as (2). In a nutshell, our solution to (P1) consists of full measure statements that are quantitative in nature. Next, turning our attention to (P2), Theorem 2 in [25] provides a precise formula for the Hausdorff dimension of $W(T, \{B_n\})$ when T is an integer matrix transformation diagonalizable over the rationals. In this paper we investigate the more general situation in which T is real and by making use of technology that was not available at the time of [25], we show (for instance) that the aforementioned formula for the Hausdorff dimension holds for a large class of real diagonal matrix transformations.

At the heart of our solution to (P1) for matrix transformations of tori, is a result that holds for a large class of measure-preserving dynamical systems. We start with describing this broader result and then move onto formally stating our theorems for matrix transformations.

1.1. *A quantitative full measure result for* Σ*-mixing dynamical systems*

Given a measure-preserving dynamical system $(X, \mathcal{B}, \mu, T)$ and a sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets in $\mathcal{B}$, we will show that if μ is exponentially mixing and the measure sum in (1) diverges then the associated lim sup set $W\big(T, \{E_n\}\big)$ is of full measure. However, it turns out that a lot more is true. We can establish a quantitative full measure statement and at the same time work with the potentially weaker notion of Σ-mixing.

Definition 1. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving dynamical system and $\mathcal{C}$ be a collection of measurable subsets of X. For $n \in \mathbb{N}$, let

$$\phi(n) := \sup\left\{ \left|\frac{\mu(E \cap T^{-n}F)}{\mu(F)} - \mu(E)\right| \;:\; E \in \mathcal{C}, F \in \mathcal{C}\right\}. \tag{3}$$

[2] The author SV would like to take this opportunity to apologise for making an announcement and then not delivering the goods!

We say that μ is Σ-mixing (short for summable-mixing) with respect to $(T, \mathcal{C})$ if the series $\sum_{n=1}^{\infty} \phi(n)$ converges.

Recall, that the above of mixing is stronger than that of *ϕ-mixing* which simply requires that $\phi(n) \to 0$ as $n \to \infty$. Also recall, that μ is *exponentially mixing* with respect to $(T, \mathcal{C})$ if there exists a constant $0 < \gamma < 1$ such that

$$\mu(E \cap T^{-n}(F)) = \mu(E)\mu(F) + O(\gamma^n)\mu(F), \tag{4}$$

for any $n \geq 1$ and $E, F \in \mathcal{C}$ – the implied constant in the big O does not depend on the sets E and F. In other words, and not surprisingly, exponentially mixing and Σ-mixing coincide whenever $\phi(n)$ converges to zero exponentially fast. It is worth mentioning that in the standard definition, condition (4) is required to hold for any $F \in \mathcal{B}$ rather than just in $\mathcal{C}$. We refer the reader to the survey paper [7] for further details including "other" variants of the notion of exponentially mixing. Also, see §1.1.1 below.

As above, let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of measurable subsets of X. Then, given $N \in \mathbb{N}$ and $x \in X$, consider the counting function

$$R(x, N) = R(x, N; T, \{E_n\}) := \#\{1 \leq n \leq N : T^n(x) \in E_n\}. \tag{5}$$

As alluded to in the definition, we will often simply write $R(x, N)$ for $R(x, N; T, \{E_n\})$ since the other dependencies will be clear from the context and are usually fixed. It is easily seen that the convergent statement (1) is equivalent to saying that if the measure sum converges, then $\lim_{N\to\infty} R(x, N)$ is finite for μ–almost all $x \in X$. The following result implies that for a large class of dynamical systems, if the measure sum diverges then μ–almost all $x \in X$ 'hit' the target sets E_n the 'expected' number of times.

Theorem 1. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving dynamical system and $\mathcal{C}$ be a collection of subsets of X. Suppose that μ is Σ-mixing with respect to $(T, \mathcal{C})$ and let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of subsets in $\mathcal{C}$. Then, for any given $\varepsilon > 0$, we have that*

$$R(x, N) = \Phi(N) + O\left(\Phi^{1/2}(N)\ (\log \Phi(N))^{3/2+\varepsilon}\right) \tag{6}$$

for μ-almost all $x \in X$, where

$$\Phi(N) := \sum_{n=1}^{N} \mu(E_n)\,.$$

A simple consequence of Theorem 1 is that $\lim_{N\to\infty} R(x, N) = \infty$ for μ–almost all $x \in X$ if the measure sum diverges and so together with (1) we obtain the following zero-full measure criterion.

Corollary 1. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving dynamical system and $\mathcal{C}$ be a collection of subsets of X. Suppose that μ is Σ-mixing with respect to $(T, \mathcal{C})$ and let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of subsets in $\mathcal{C}$. Then*

$$\mu\big(W(T, \{E_n\})\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} \mu\big(E_n\big) < \infty \\ 1 & \text{if } \sum_{n=1}^{\infty} \mu\big(E_n\big) = \infty. \end{cases} \tag{7}$$

Before moving onto considering the specific situation in which T is a matrix transformation of the torus we discuss previous related works.

1.1.1. Connection to other works

We will focus on two previous works that are related to the framework presented above; i.e. the notion of Σ-mixing and its consequences. In an interesting paper [19], Fernández, Melián & Pestana introduced the notion of a transformation T being *uniformly mixing at a point* $x_0 \in X$. Their notion coincides with our Definition 1 if we restrict the collection $\mathcal{C}$ to balls B centered at x_0. The upshot [19, Theorem 1] is that given a decreasing sequence of balls $B_n := B(x_0, r_n)$, if T is uniformly mixing at x_0 and $\sum_{n=1}^{\infty} \mu(B_n) = \infty$ then

$$\lim_{N\to\infty} \frac{R(x,N)}{\Phi(N)} = \lim_{N\to\infty} \frac{\#\big\{1 \le n \le N : T^n(x) \in B_n\big\}}{\sum_{n=1}^{N} \mu(B_n)} = 1\,. \tag{8}$$

Clearly our Theorem 1 not only implies this asymptotic statement but it also provides a reasonably sharp estimate for the error term. As a consequence, the various applications of (8) considered in [19] can be strengthened accordingly. Indeed, their main motivating application to inner functions [19, Theorem 2] can be improved to the following statement.

Theorem FMP$^+$. *Let $f : \mathbb{D} \to \mathbb{D}$ be an inner function with $f(0) = 0$, but not a rotation. Let ξ_0 be a point in $\partial\mathbb{D}$ and let $\{r_n\}$ be a decreasing sequence of positive numbers. If $\sum_{n=1}^{\infty} r_n = \infty$, then for any given $\varepsilon > 0$, we have that*

$$\#\big\{1 \le n \le N : d\big((f^*)^n(\xi), \xi_0\big) < r_n\big\} = \Phi(N) + O\left(\Phi^{1/2}(N)\ (\log\Phi(N))^{3/2+\varepsilon}\right)$$

for μ-almost all $x \in X$, where $\Phi(N) := \sum_{n=1}^{N} r_n$, $f^*(\xi) = \lim_{r\to 1_-} f(r\xi)$ and d is the angular distance in $\partial\mathbb{D}$.

In the later stages of preparing this manuscript, we discovered that our Theorem 1 overlaps with a result of Philipp [42, Theorem 3] dating back to 1967. Indeed, in his theorem the condition imposed on the sequence $\{E_n\}_{n\in\mathbb{N}}$ of measurable sets is in effect equivalent to our notion of Σ-mixing with $\mathcal{C} = \mathcal{B}$. It appears that [42, Theorem 3] has been either entirely overlooked, or at least not fully exploited in previous works. For the sake of completeness we have decided to include the proof of Theorem 1 in §2. Moreover, our proof is pretty short and unlike Philipp's approach it exploits a rather general tool (Lemma 1 in §2) for establishing sharp counting statements. To the best of our knowledge, a slightly weaker version of the tool, which suffices to establish Theorem 1, first appears in Sprindžuk's book [45, Lemma 10] which was some ten years after Philipp's paper. Furthermore, we have decided to include a self contained proof of the Corollary 1 since it is rather nifty and some readers may only be interested in the zero-full measure criterion rather than its stronger quantitative form.

1.2. Quantitative full measure results for matrix transformations

Let T be a $d \times d$ non-singular matrix with real coefficients. Then, T determines a self-map of the d-dimensional torus $X = \mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$; namely, it sends $\mathbf{x} \in \mathbb{T}^d$ to $T\mathbf{x}$ modulo one. In what follows, T will denote both the matrix and the transformation. It should be obvious from the context what is meant. Furthermore, for $n \in \mathbb{N}$, by T^n we will always mean the n-th iteration of the transformation T rather than the matrix multiplied n times. With reference to the general setup of §1.1, we now describe a broad collection $\mathcal{C}$ of 'target' sets contained in $\mathbb{T}^d$ so that for any sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets in $\mathcal{C}$ we are able to address the shrinking target 'measure' problem (P1) for the associated lim sup set $W\big(T, \{E_n\}\big)$. In order to do this, we require the notion of the Minkowski content of a set in $\mathbb{R}^d$. We start by recalling this basic notion from geometric measure theory.

Let $0 \le s \le d$ be two positive integers and let A be a subset of $\mathbb{R}^d$. Let m_d denote the d-dimensional Lebesgue measure and $\alpha(d)$ denote the volume of the d-dimensional unit Euclidean ball $\{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| < 1\}$. By convention, we define $\alpha(0) := 1$. For $0 < \epsilon < \infty$, we let $A(\epsilon)$ denote the ϵ-neighbourhood of A; that is

$$A(\epsilon) := \{\mathbf{x} \in \mathbb{R}^d : \mathrm{dist}(\mathbf{x}, A) < \epsilon\}\,.$$

Then, following the classical text of Federer [18, Section 3.2.37], the s-dimensional upper and lower Minkowski content of A are defined, respectively as

$$M^{*s}(A) := \limsup_{\epsilon\to 0^+} \frac{m_d(A(\epsilon))}{\alpha(d-s)\,\epsilon^{d-s}} \qquad \text{and} \qquad M_*^s(A) := \liminf_{\epsilon\to 0^+} \frac{m_d(A(\epsilon))}{\alpha(d-s)\epsilon^{d-s}}\,.$$

If these upper and lower Minkowski contents are equal, then their common value is called the s-dimensional Minkowski content of A and is denoted by $M^s(A)$. In general, the set functions M^{*s} and M_*^s are not measures. However, for nice sets it turns out that both equal a constant multiple of the Lebesgue measure m_s. In particular, a result of Federer [18, Theorem 3.2.39] states that if A is a closed s-rectifiable subset of $\mathbb{R}^d$ (i.e. the image of a bounded set from $\mathbb{R}^s$ under a Lipschitz function), then the s-dimensional Minkowski content of A exists, and is equal to the s-dimensional Hausdorff measure of A. Recall that for integer s the latter is a constant multiple of s-dimensional Lebesgue measure. Also, for the sake of completeness it is worth mentioning that the Minkowski content is intimately related to the Minkowski dimension which, nowadays is more commonly referred to as the box dimension. When considering this fractal dimension, s need not be an integer and we put $\alpha(d-s)=1$ in the above definitions of upper and lower Minkowski contents. For further details see [14, Section 3.1], [18, Sections 3.2.37-44], [35, Chapter 5] and references within.

The following proposition identifies the collection $\mathcal{C}$ of 'target' sets alluded to above as subsets E of $\mathbb{T}^d$ for which the boundary ∂E has bounded $(d-1)$-dimensional upper Minkowski content. It makes use of the work initiated by Keller [29,30] on the existence and properties of absolutely continuous invariant measures for piecewise expanding maps, and subsequently developed by the likes of Góra & Boyarsky [22], Buzzi [9–11], Buzzi & Maume-Deschamps [12], Saussol [43] and Tsujii [47,48].

Proposition 1. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that all eigenvalues of T are of modulus strictly larger than 1. Then*

- *(i) there exists an absolutely continuous (with respect to Lebesgue measure m_d) invariant probability measure (acim) μ,*
- *(ii) the support $A \subseteq \mathbb{T}^d$ of any acim μ can be decomposed into finitely many disjoint measurable sets $A_1, \ldots, A_s$ such that for each $1 \leq i \leq s$ the restriction $\mu|_{A_i}$ of μ to A_i is ergodic and is equivalent to the restriction $m_d|_{A_i}$ of Lebesgue measure m_d to A_i,*
- *(iii) each ergodic component A_i in (ii) can in turn be decomposed into finitely many disjoint measurable sets $A_{i1}, \ldots, A_{ip_i}$ such that for each $1 \leq j \leq p_i$ the restriction $\mu|_{A_{ij}}$ is mixing with respect to T^{p_i},*
- *(iv) on each mixing component A_{ij} in (iii), the restriction $\mu|_{A_{ij}}$ is exponentially mixing with respect to $(T^{p_i}, \mathcal{C})$ for any collection $\mathcal{C}$ of subsets E of A_{ij} satisfying the bounded property*

$$(\mathbf{B}): \quad \sup_{E \in \mathcal{C}} M^{*(d-1)}(\partial E) < \infty .$$

Remark 1. By definition, the restriction $\mu|_A$ of a probability measure μ to a set A with $\mu(A) > 0$ is normalized so that it too is a probability measure. In other words, for an arbitrary measurable set E

$$\mu|_A(E) := \frac{1}{\mu(A)}\, \mu(E) .$$

For each $1 \leq i \leq s$, the sets A_{ij} $(1 \leq j \leq p_i)$ appearing in part (iii) are referred to as the mixing components of A_i $(= \bigcup_{j=1}^{p_i} A_{ij})$ and the positive integers p_i are the period of the mixing components. These mixing components satisfy the property that

$$T(A_{ij}) = A_{ij+1} \quad (1 \leq j \leq p_i - 1) \qquad \text{and} \qquad T(A_{ip_i}) = A_{i1} .$$

Also, for the sake of clarity, completeness and convenience, recall that if μ and ν are two measures on the same measurable space, then μ is *absolutely continuous* with respect to ν (written $\mu \ll \nu$) if $\mu(E) = 0$ for every measurable set E for which $\nu(E) = 0$. Moreover, the measures μ and ν are *equivalent* if $\mu \ll \nu$ and $\nu \ll \mu$ and are said to be *strongly equivalent* or *comparable* if there exists a constant $C \geq 1$ such that for every measurable set E

$$C^{-1}\nu(E) \leq \mu(E) \leq C\nu(E).$$

Given a measure-preserving dynamical system $(X, \mathcal{B}, \mu, T)$, the invariant measure μ is ergodic if for every set $E \in \mathcal{B}$ with $T^{-1}E = E$ we have either $\mu(E) = 0$ or $\mu(E) = 1$. Moreover, μ is said to be mixing with respect to T (often referred to strong-mixing) if for every $E, F \in \mathcal{B}$

$$\lim_{n\to\infty} \mu(E \cap T^{-n}F) = \mu(E)\mu(F).$$

Clearly, exponentially mixing tells us that the implied error term in the above limit decays exponentially. Also, if we put $F = E$ we immediately see that mixing implies ergodic.

The following constitutes our most general measure theoretic result for the shrinking target problem for matrix transformations of tori. As we shall see the "divergent" part, which is the hard part, is essentially an immediate consequence of combining Theorem 1 and Proposition 1.

Theorem 2. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that all eigenvalues of T are of modulus strictly larger than 1 and let $\mathcal{C}$ be any collection of subsets E of $\mathbb{T}^d$ satisfying the bounded property* (**B**). *Furthermore, let μ be an acim and suppose it has support $\mathbb{T}^d$ and is mixing with respect to T. Then for any sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets in $\mathcal{C}$ and $\varepsilon > 0$, we have that*

$$R(\mathbf{x}, N) = \Phi(N) + O\left(\Phi^{1/2}(N)\ (\log \Phi(N))^{3/2+\varepsilon}\right) \tag{9}$$

for μ-almost all (equivalently m_d-almost all) $\mathbf{x} \in \mathbb{T}^d$, where $\Phi(N) := \sum_{n=1}^{N} \mu(E_n)$. In particular,

$$m_d\big(W(T, \{E_n\})\big) = \mu\big(W(T, \{E_n\})\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} \mu\big(E_n\big) < \infty \\ \\ 1 & \text{if } \sum_{n=1}^{\infty} \mu\big(E_n\big) = \infty. \end{cases} \tag{10}$$

We note that the existence of the acim measure μ in Theorem 2 is guaranteed by part (i) of Proposition 1 and that the assumptions imposed on it, namely that the support of μ

is the whole space $\mathbb{T}^d$ and that μ is mixing with respect to T, are often satisfied. Indeed, this is the situation when the eigenvalues of T are large in modulus or the coefficients of T are integers. Regarding the former we have the following precise statement. We will come to the integer situation shortly (see Theorem 5 below).

Theorem 3. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that all eigenvalues of T are of modulus strictly larger than $1+\sqrt{d}$. Let $\mathcal{C}$ be any collection of subsets of $\mathbb{T}^d$ satisfying the bounded property* (**B**)*. Then there is a unique acim μ, such that for any sequence $\{E_n\}_{n\in\mathbb{N}}$ in $\mathcal{C}$ and $\varepsilon > 0$, the counting formula* (9) *holds for μ-almost all (equivalently m_d-almost all) $\mathbf{x} \in \mathbb{T}^d$, where $\Phi(N) := \sum_{n=1}^{N} \mu(E_n)$. In particular, the zero-full measure criteria* (10) *holds.*

Remark 2. The fact that the acim μ appearing in Theorem 3 is unique is a trivial consequence of the fact that any acim satisfying the hypotheses of Theorem 2 has to be unique. Indeed, to see that this is the case, suppose there exist two such measures. Then by part (ii) of Proposition 1, both are equivalent to m_d. By assumption, both are mixing with respect to T and hence ergodic. It thus follows (see [51, Theorem 6.10]) that the two measures are equal.

Remark 3. By using the full force of Proposition 1, the assumptions on μ in Theorem 2 can be completely dropped if we restrict our attention to the shrinking target set $W(T^p, \{E_n\}) \cap A$. Here $A \subseteq \mathbb{T}^d$ is the support of the acim μ (guaranteed by part (i) of Proposition 1) and $p := p_1 p_2 \dots p_s$ where the integers p_i are the periods of the mixing components associated with part (iii) of Proposition 1. Establishing Theorem 4 below is an illustration of precisely this remark in action. In short, the point of making the assumptions on μ in Theorem 2 is to obtain a simple statement for the size of $W(T, \{E_n\})$ in terms of the probability measure m_d supported on $\mathbb{T}^d$.

Remark 4. We consider two special families of target sets that correspond to "natural" setups within the classical theory of Diophantine approximation. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive function and fix some point $\mathbf{a} := (a_1, \dots, a_d) \in \mathbb{T}^d$. For $n \in \mathbb{N}$, let

$$B_n = B(\mathbf{a}, \psi(n)) := \left\{ \mathbf{x} \in \mathbb{T}^d : \max_{1\le i\le d} \|x_i - a_i\| \le \psi(n) \right\}$$

and

$$H_n = H(\mathbf{a}, \psi(n)) := \left\{ \mathbf{x} \in \mathbb{T}^d : \prod_{1\le i\le d} \|x_i - a_i\| \le \psi(n) \right\}, \tag{11}$$

where $\| \,.\, \|$ denotes the distance to the nearest integer. Clearly, B_n is a ball with respect to the maximum norm and H_n is a hyperboloid – both are centred at the fixed point $\mathbf{a}$. In turn, let

$$W(T,\psi,\mathbf{a}) := \{\mathbf{x} \in \mathbb{T}^d : T^n(\mathbf{x}) \in B(\mathbf{a},\psi(n)) \text{ for infinitely many } n \in \mathbb{N}\},$$

and

$$W^\times(T,\psi,\mathbf{a}) := \{\mathbf{x} \in \mathbb{T}^d : T^n(\mathbf{x}) \in H(\mathbf{a},\psi(n)) \text{ for infinitely many } n \in \mathbb{N}\},$$

denote the corresponding shrinking target sets. The former is intimately related to sets studied within the classical simultaneous theory of Diophantine approximation and the latter to the multiplicative theory. To see this explicitly, suppose that T is an integer, diagonal matrix. In fact, suppose that

$$T = \operatorname{diag}(t_1,\ldots,t_d) \quad \text{with} \quad t_i \geq 2$$

and for convenience suppose $\mathbf{a}$ is the origin $\mathbf{0}$. Then, on using the fact that T is integer, it is easily seen that for any given $\mathbf{x} := (x_1,\ldots,x_d) \in [0,1)^d$ we have

$$T^n(\mathbf{x}) \in B(\mathbf{0},\psi(n)) \quad \Longleftrightarrow \quad \max_{1\leq i\leq d} \|t_i^n x_i\| \leq \psi(n)$$

and

$$T^n(\mathbf{x}) \in H(\mathbf{0},\psi(n)) \quad \Longleftrightarrow \quad \prod_{1\leq i\leq d} \|t_i^n x_i\| \leq \psi(n)\,.$$

It is evident that both the families of target sets $\{B_n\}_{n\geq 1}$ and $\{H_n\}_{n\geq 1}$ satisfy the bounded property (**B**). Thus, at the very least, our theorems incorporate both the simultaneous and multiplicative aspects of the classical theory of Diophantine approximation in which the denominators of the rational approximates are restricted to lacunary sequences. For the explicit statements see Corollaries 2 & 3 below. In fact, our bounded property (**B**) condition is far more general than the so called property (**P**) condition (see §5.1) imposed by Gallagher in his elegant and influential paper [20]. We reiterate that our results hold for any family of target sets $\{E_n\}_{n\geq 1}$ whose boundaries are rectifiable and are of uniformly bounded $(d-1)$-dimensional Lebesgue measure.

We now investigate natural situations in which the measure μ associated with Theorem 2 is strongly equivalent to Lebesgue measure m_d on $\mathbb{T}^d$. For such situations we can replace μ by m_d in the finite sum $\Phi(N)$ and the righthand side of (10) and thus obtain statements entirely in terms of Lebesgue measure. To start with, let us stick with real, non-singular matrices and suppose that T is diagonal with all eigenvalues (or equivalently diagonal entries) $\beta_1,\beta_2,\ldots,\beta_d$ of modulus strictly larger than 1. Now with this in mind, let $\beta \in \mathbb{R}$ such that $|\beta| > 1$ and let μ_β be corresponding Parry measure for positive β or the Yrrap measure for negative β – see §3.3 for background and further details. Also, let $K(\beta)$ denote the support of μ_β. Then (see Proposition 3 below),

$$K(\beta) = [0,1] \qquad \text{if} \qquad \beta \in (-\infty, -g] \cup (1, +\infty)\,, \tag{12}$$

and $K(\beta)$ is a finite union of closed intervals contained in $[0,1]$ if $\beta \in (-g, -1)$. Here and throughout,

$$g := (\sqrt{5}+1)/2$$

is the golden ratio. Now returning to the transformation T of the torus $\mathbb{T}^d$, we consider the product measure ν of the corresponding one-dimensional Parry-Yrrap measures μ_{β_i}; that is

$$\nu := \mu_{\beta_1} \times \mu_{\beta_2} \times \cdots \times \mu_{\beta_d}\,. \tag{13}$$

Then by definition, the support of ν is

$$K := \prod_{i=1}^{d} K(\beta_i),$$

and in view of (12) we have that $K = \mathbb{T}^d$ if all β_i are in $(-\infty, -g] \cup (1, +\infty)$. On exploiting the properties of the Parry-Yrrap measures μ_{β_i} and using the full force of Proposition 1 (see Remark 3) we are able to show that ν is exponentially mixing with respect to $(T, \mathcal{C})$ for any collection $\mathcal{C}$ of subsets E of K satisfying the bounded property (**B**). The details of this are given in §3.3 and is at the heart of establishing the following statement for real, diagonal matrix transformations.

Theorem 4. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal and all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ are of modulus strictly larger than 1. Let ν be the product measure given by (13) with support $K \subseteq \mathbb{T}^d$. Let $\mathcal{C}$ be any collection of subsets E of K satisfying the bounded property (**B**). Then for any sequence $\{E_n\}_{n\in\mathbb{N}}$ in $\mathcal{C}$ and $\varepsilon > 0$, the counting formula (9) holds for ν-almost all (equivalently $m_d|_K$-almost all) $\mathbf{x} \in \mathbb{T}^d$, where $\Phi(N) := \sum_{n=1}^{N} \nu(E_n)$. In particular,*

$$m_d|_K\big(W(T, \{E_n\})\big) = \nu\big(W(T, \{E_n\})\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} \nu\big(E_n\big) < \infty \\[2ex] 1 & \text{if } \sum_{n=1}^{\infty} \nu\big(E_n\big) = \infty. \end{cases}$$

Furthermore, if all the eigenvalues of T are in $(-\infty, -g] \cup (1, +\infty)$ then $K = \mathbb{T}^d$ and we can replace ν by m_d in the above; i.e. the counting formula (9) holds for m_d-almost all $\mathbf{x} \in \mathbb{T}^d$, where $\Phi(N) := \sum_{n=1}^{N} m_d(E_n)$ and in particular

$$m_d\big(W(T, \{E_n\})\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} m_d\big(E_n\big) < \infty \\[2ex] 1 & \text{if } \sum_{n=1}^{\infty} m_d\big(E_n\big) = \infty. \end{cases}$$

In the case the collection $\mathcal{C}$ of subsets of K is restricted to rectangles with sides parallel to the axes (they clearly satisfy the bounded property (**B**)) we can avoid using Proposition 1 and give a self-contained and reasonably elementary proof of the above theorem (see §3.3.1). In particular, it is more than enough to establish the following corollary for balls (cubes); i.e., when we take $E_n = B_n$ (see Remark 4) in the above theorem.

Corollary 2. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal with all eigenvalues in $(-\infty, -g] \cup (1, +\infty)$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be real positive function and $\mathbf{a} \in \mathbb{T}^d$. Then for any $\varepsilon > 0$, we have that*

$$\#\left\{1 \le n \le N : T^n(\mathbf{x}) \in B(\mathbf{a}, \psi(n))\right\} = \Phi(N) + O\left(\Phi^{1/2}(N)\ (\log \Phi(N))^{3/2+\varepsilon}\right)$$

for m_d-almost all $\mathbf{x} \in \mathbb{T}^d$, where $\Phi(N) := \sum_{n=1}^{N} (2\psi(n))^d$. In particular,

$$m_d\big(W(T, \psi, \mathbf{a})\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} \psi(n)^d < \infty \\ 1 & \text{if } \sum_{n=1}^{\infty} \psi(n)^d = \infty. \end{cases}$$

In fact, if we assume that $\psi(n) \to 0$ as $n \to \infty$, we are able to appropriately extend Corollary 2 to the situation in which the eigenvalues are in $(-\infty, -g] \cup [-1, +\infty)$. In other words, we can incorporate the interval [-1,1] into the allowed range of the eigenvalues. This is the subject of §3.3.2 below.

In another direction, if T is an integer matrix transformation we are able to use a nifty "reduction" argument to relax the condition that T is diagonal in Theorem 4 to T is diagonalizable over $\mathbb{Z}$. This reduction argument is the subject of §3.3.3 below. In fact, for integer matrices far more is true.

Theorem 5. *Let T be an integer, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that all eigenvalues of T are of modulus strictly larger than 1 and let $\mathcal{C}$ be any collection of subsets E of $\mathbb{T}^d$ satisfying the bounded property (**B**). Then, for any sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets in $\mathcal{C}$ and $\varepsilon > 0$, the counting formula (9) holds for m_d-almost all $\mathbf{x} \in \mathbb{T}^d$, where $\Phi(N) := \sum_{n=1}^{N} m_d(E_n)$. In particular,*

$$m_d\big(W(T, \{E_n\})\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} m_d(E_n) < \infty \\ 1 & \text{if } \sum_{n=1}^{\infty} m_d(E_n) = \infty. \end{cases}$$

To end with, we illustrate natural "number theoretic" consequences of our results. Let $t_1, \ldots, t_d \ge 2$ be integers and let $T = \operatorname{diag}(t_1, \ldots, t_d)$. Then with reference to Remark 4, it follows that

$$W(T, \psi, \mathbf{a}) = \{\mathbf{x} \in [0,1)^d : \max_{1\le i\le d} \|t_i^n x_i - a_i\| \le \psi(n) \ \text{ for infinitely many } n \in \mathbb{N}\}$$

and

$$W^{\times}(T,\psi,\mathbf{a})=\{\mathbf{x}\in[0,1)^d:\prod_{1\le i\le d}\|t_i^n x_i-a_i\|\le\psi(n)\ \text{ for infinitely many } n\in\mathbb{N}\}.$$

Thus, Theorem 5 implies the following statement for multiplicative Diophantine approximation. In fact, since T is diagonal, it is also covered by the "furthermore part" of Theorem 4.

Corollary 3. *Let $t_1,\ldots,t_d\ge 2$ be integers and let $T=\operatorname{diag}(t_1,\ldots,t_d)$. Let $\psi:\mathbb{R}^+\to\mathbb{R}^+$ be real positive function such that $\psi(x)<2^{-d}$ and $\mathbf{a}=(a_1,\ldots,a_d)\in\mathbb{T}^d$. Then*

$$\#\Big\{1\le n\le N:\prod_{1\le i\le d}\|t_i^n x_i-a_i\|\le\psi(n)\Big\}=\Phi(N)+O\left(\Phi^{1/2}(N)\ (\log\Phi(N))^{3/2+\varepsilon}\right)$$

for m_d-almost all $\mathbf{x}=(x_1,\ldots,x_d)\in\mathbb{T}^d$, where

$$\Phi(N)=\sum_{n=1}^{N}2^d\psi(n)\left(\sum_{s=0}^{d-1}\frac{1}{s!}\left(\log\frac{1}{2^d\psi(n)}\right)^s\right)$$

In particular,

$$m_d\big(W^{\times}(T,\psi,\mathbf{a})\big)=\begin{cases}0 & \text{if } \sum_{n=1}^{\infty}\psi(n)\left(\log\frac{1}{\psi(n)}\right)^{d-1}<\infty\\[2ex] 1 & \text{if } \sum_{n=1}^{\infty}\psi(n)\left(\log\frac{1}{\psi(n)}\right)^{d-1}=\infty.\end{cases}$$

The analogous statement for the simultaneous set $W(T,\psi,\mathbf{a})$ is clearly covered by Corollary 2 above. The condition that $\psi(x)<2^{-d}$ is only required for the counting statement.

Corollary 3 is probably most familiar to number theorists within the context of when $t_1=\ldots=t_d$. This corresponds to approximating arbitrary points $\mathbf{x}\in[0,1]^d$ by "shifted" rational points $((p_1+a_1)/q,\ldots,(p_d+a_d)/q)$ with denominators q restricted to an integer lacunary sequence. In this setup, the zero-full measure criterion within the corollary can just as easily be deduced from the elegant work of Gallagher [20] mentioned in Remark 4. Also, under the same setup and the assumption that ψ is non-increasing, the corresponding quantitative version (with a slightly worse error term) can be deduced from [23, Theorem 4.6].

Remark 5. For $n\in\mathbb{N}$, let $H_n=H(\mathbf{a},\psi(n))$ be the hyperboloid region given by (11). Clearly, Corollary 3 follows directly from Theorem 5 on letting $E_n=H_n$ and on showing that

$$m_d(H_n)=2^d\psi(n)\left(\sum_{s=0}^{d-1}\frac{1}{s!}\left(\log\frac{1}{2^d\psi(n)}\right)^s\right). \tag{14}$$

For the sake of completeness we will provide the details of this measure calculation in §3.2.1.

1.3. Dimension results for matrix transformations

We address the shrinking target 'dimension' problem (P2) in the case T is a self-map of the d-dimensional torus $\mathbb{T}^d$ and the target sets are a sequence $\{B_n\}_{n\in\mathbb{N}}$ of balls as in the original formulation of the problem. The following two theorems constitute our main dimension results. It turns out that these statements for balls can be exploited to determine the dimension of shrinking targets sets in the case the targets are a sequence $\{H_n\}_{n\in\mathbb{N}}$ of hyperboloids. Throughout, given a real positive function $\psi:\mathbb{R}^+\to\mathbb{R}^+$ we let $\lambda=\lambda(\psi)$ denote its lower order at infinity; that is

$$\lambda=\lambda(\psi):=\liminf_{n\to\infty}\frac{-\log\psi(n)}{n}.$$

Theorem 6. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal with all eigenvalues $\beta_1,\beta_2,\ldots,\beta_d$ strictly larger than 1. Assume that $1<\beta_1\le\beta_2\le\cdots\le\beta_d$. Let $\psi:\mathbb{R}^+\to\mathbb{R}^+$ be a real positive function and $\mathbf{a}\in\mathbb{T}^d$. Then*

$$\dim_{\mathrm{H}} W(T,\psi,\mathbf{a})=\min_{1\le i\le d}\theta_i(\lambda),$$

where

$$\theta_i(\lambda):=\frac{i\log\beta_i-\sum\limits_{k:\beta_k>\beta_i e^{\lambda}}(\log\beta_j-\log\beta_i-\lambda)+\sum\limits_{k>i}\log\beta_j}{\lambda+\log\beta_i}.$$

Remark 6. We will in fact deduce the above theorem from a more general statement concerning rectangular target sets – see Theorem 12 in §4.2.

In the case $d=1$, the above result corresponds to the main result in [44]. It turns out that while we are currently unable to prove in full generality the analogue of Theorem 6 that incorporates negative eigenvalues, we can do so in the one dimensional case. Thus, the following statement for $\beta<-1$ is new and extends the work of Shen & Wang [44] from positive to arbitrary β-transformations T_β.

Theorem 7. *Let β be a real number with $|\beta|>1$ and $K(\beta)$ be the support of the associated Parry-Yrrap measure. Let $\psi:\mathbb{R}^+\to\mathbb{R}^+$ be a real positive decreasing function and $a\in K(\beta)$. Then*

$$\dim_{\mathrm{H}} W(T_\beta,\psi,a)=\frac{\log|\beta|}{\lambda+\log|\beta|}.$$

The proof of Theorem 7 makes use of a general approximation technique for any piecewise linear map of the unit interval with constant slope. The associated result (Proposition 7 in §7) may prove to be useful for other problems.

We now mention two consequences of our main dimension theorems. The first is that if T is an integer matrix transformation, then in Theorem 6 we can replace the condition that T is diagonal by T is diagonalizable over $\mathbb{Z}$.

Theorem 8. *Let T be an integer, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonalizable over $\mathbb{Z}$ with all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ strictly larger than 1. Assume that $1 < \beta_1 \le \beta_2 \le \cdots \le \beta_d$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive function and $\mathbf{a} \in \mathbb{T}^d$. Then*

$$\dim_{\mathrm{H}} W(T, \psi, \mathbf{a}) = \min_{1 \le i \le d} \theta_i(\lambda) .$$

The theorem follows from Theorem 6 by using a "reduction" argument – see §4.2.2. The second is that the above theorems for balls enables us to establish the dimension of the multiplicative set $W^\times(T, \psi, \mathbf{a})$. In fact, we only require the $d = 1$ statement and so we are able to utilize the more general Theorem 7.

Theorem 9. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal and all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ are of modulus strictly larger than 1. Assume that $1 < |\beta_1| \le |\beta_2| \le \cdots \le |\beta_d|$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive decreasing function and $\mathbf{a} \in \mathbb{T}^d$ with $a_d \in K(\beta_d)$. Then*

$$\dim_{\mathrm{H}} W^\times(T, \psi, \mathbf{a}) = d - 1 + \frac{\log|\beta_d|}{\lambda + \log|\beta_d|} .$$

This consequence of Theorem 7 was pointed out to us by Baowei Wang. We thank him for sharing his insight and indeed for providing the details of the proof which forms the appendix. We stress that making use of Theorem 7, rather than the previously known $d = 1$ case of Theorem 6 due to Shen & Wang [44] is crucial. The latter requires that all the eigenvalues are positive and strictly larger than one and would thus yield a weaker version of Theorem 9.

2. Establishing Theorem 1 and Corollary 1

The following statement [23, Lemma 1.5] represents an important tool in the theory of metric Diophantine approximation for establishing counting statements. It has its bases in the familiar variance method of probability theory and can be viewed as the quantitative form of the (divergence) Borel-Cantelli Lemma [2, Lemma 2.2].

Lemma 1. *Let $(X, \mathcal{B}, \mu)$ be a probability space, let $(f_n(x))_{n\in\mathbb{N}}$ be a sequence of non-negative μ-measurable functions defined on X, and $(f_n)_{n\in\mathbb{N}}$, $(\phi_n)_{n\in\mathbb{N}}$ be sequences of real numbers such that*

$$0 \le f_n \le \phi_n \quad (n = 1, 2, \ldots).$$

Suppose that for arbitrary $a, b \in \mathbb{N}$ with $a < b$, we have

$$\int_X \left(\sum_{n=a}^{b} \big(f_n(x) - f_n \big) \right)^2 \mathrm{d}\mu(x) \;\le\; C \sum_{n=a}^{b} \phi_n \tag{15}$$

for an absolute constant $C > 0$. Then, for any given $\varepsilon > 0$, we have

$$\sum_{n=1}^{N} f_n(x) \;=\; \sum_{n=1}^{N} f_n \;+\; O\left(\Phi(N)^{1/2} \log^{\frac{3}{2}+\varepsilon} \Phi(N) + \max_{1\le k\le N} f_k \right) \tag{16}$$

for μ-almost all $x \in X$, where $\Phi(N) := \sum\limits_{n=1}^{N} \phi_n$.

Note that in statistical terms, if the sequence f_n is the mean of $f_n(x)$; i.e.

$$f_n = \int_X f_n(x) \mathrm{d}\mu(x)\,,$$

then the l.h.s. of (15) is simply the variance $\mathrm{Var}(Z_{a,b})$ of the random variable

$$Z_{a,b} = Z_{a,b}(x) := \sum_{n=a}^{b} f_n(x)\,.$$

In particular,

$$\mathrm{Var}(Z_{a,b}) = \mathbb{E}(Z_{a,b}^2) - \mathbb{E}(Z_{a,b})^2$$

where

$$\mathbb{E}(Z_{a,b}) = \int_X Z_{a,b}(x) \mathrm{d}\mu(x)\,.$$

The following extremely useful classical inequality that bounds the probability that a random variable is small, in terms of its expectation and second moment, is a well know consequence of the Cauchy-Schwarz inequality.

Lemma 2 *(Paley-Zygmund Inequality). Let $(X, \mathcal{B}, \mu)$ be a probability space and Z be a non-negative random variable. Then for any $0 < \lambda < 1$, we have that*

$$\mu\Big(\{x \in X : Z(x) > \lambda \mathbb{E}(Z)\}\Big) \geq (1-\lambda)^2 \, \frac{\mathbb{E}(Z)^2}{\mathbb{E}(Z^2)} .$$

We shall see that a straightforward application of Lemma 2 leads to a direct proof of Corollary 1. Deducing Theorem 1 from Lemma 1 is also pretty straightforward.

Proof of Theorem 1. Given a sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets in $\mathcal{C}$, we consider Lemma 1 with

$$f_n(x) := \chi_{T^{-n}(E_n)}(x) = \chi_{E_n}(T^n(x)) \quad \text{and} \quad f_n := \phi_n := \mu(E_n) , \tag{17}$$

where χ_E is the characteristic function of the set $E \subset X$. Then, clearly for any $x \in X$ and $N \in \mathbb{N}$, we have that the

$$\text{l.h.s. of (16)} \ = \ \#\big\{1 \leq n \leq N : T^n(x) \in E_n\big\} := \ R(x, N)$$

and so (16) and (6) coincide. Thus, to complete the proof of Theorem 1 we need to verify that (15) is satisfied. Note that by definition, f_n is the mean of $f_n(x)$ and so

$$\text{l.h.s. of (15)} \ = \ \mathrm{Var}(Z_{a,b}) \ = \ \mathbb{E}(Z_{a,b}^2) - \mathbb{E}(Z_{a,b})^2 \tag{18}$$

where

- $\mathrm{Var}(Z_{a,b})$ is variance of the random variable

$$Z_{a,b} = Z_{a,b}(x) = \sum_{n=a}^{b} \chi_{E_n}(T^n(x)) = \sum_{n=a}^{b} \chi_{T^{-n}E_n}(x), \tag{19}$$

- the expectation

$$\mathbb{E}(Z_{a,b}) = \sum_{n=a}^{b} \mu(E_n) , \tag{20}$$

- and the second moment

$$\begin{aligned} \mathbb{E}(Z_{a,b}^2) &= \sum_{a\leq m,n\leq b} \mu\left(T^{-m}(E_m) \cap T^{-n}(E_n)\right) \\ &= \sum_{a\leq n\leq b} \mu(E_n) + 2 \sum_{a\leq m<n\leq b} \mu\left(E_m \cap T^{-(n-m)}(E_n)\right) . \end{aligned} \tag{21}$$

By making use of the Σ-mixing property (3), it follows that

$$\begin{aligned}
&\sum_{a\leq m<n\leq b} \mu\left(E_m \cap T^{-(n-m)}(E_n)\right) \\
&\leq \sum_{a\leq m<n\leq b} \mu(E_m)\mu(E_n) \;+\; \sum_{a\leq m<n\leq b} \phi(n-m)\;\mu(E_n) \\
&\leq \sum_{a\leq m<n\leq b} \mu(E_m)\mu(E_n) \;+\; \sum_{a\leq n\leq b} \left(\sum_{a\leq m<n} \phi(n-m)\right) \mu(E_n) \\
&\leq \sum_{a\leq m<n\leq b} \mu(E_m)\mu(E_n) \;+\; \kappa \sum_{a\leq n\leq b} \mu(E_n)
\end{aligned}$$

where $\kappa := \sum_{n=1}^{\infty} \phi(n) < \infty$. This together with (21) implies that

$$\begin{aligned}
\mathbb{E}(Z_{a,b}^2) &\leq (2\kappa+1) \sum_{a\leq n\leq b} \mu(E_n) \;+\; 2 \sum_{a\leq m<n\leq b} \mu(E_m)\mu(E_n) \\
&\leq (2\kappa+1) \sum_{a\leq n\leq b} \mu(E_n) \;+\; \left(\sum_{a\leq n\leq b} \mu(E_n)\right)^2 . \qquad (22)
\end{aligned}$$

The upshot of (18), (20) and (22) is that

$$\mathrm{Var}(Z_{a,b}) \;\leq\; (2\kappa+1) \sum_{a\leq n\leq b} \mu(E_n)\,.$$

This verifies (15) with $C = 2\kappa + 1$ and thereby completes the proof of Theorem 1. □

Proof of Corollary 1. In view of (1), we assume that the sum in (7) diverges. With the same notation as in the proof of Theorem 1, we start with the observation that for any $\lambda > 0$

$$\mu\Big(W(T,\{E_n\})\Big) \;\geq\; \mu\Big(\limsup_{b\to\infty}(Z_{1,b} > \lambda\mathbb{E}(Z_{1,b}))\Big) \;\geq\; \limsup_{b\to\infty} \mu\Big(Z_{1,b} > \lambda\mathbb{E}(Z_{1,b})\Big)\,. \qquad (23)$$

To estimate the measure on the r.h.s. we use the Paley-Zygmund inequality (Lemma 2) and the estimates (20) and (22). With this in mind, for any $0 < \lambda < 1$, it follows that

$$\begin{aligned}
\mu\Big(Z_{1,b} > \lambda\mathbb{E}(Z_{1,b})\Big) &\geq (1-\lambda)^2 \;\frac{\mathbb{E}(Z_{1,b})^2}{\mathbb{E}(Z_{1,b}^2)} \\
&\geq (1-\lambda)^2 \;\frac{\Big(\sum\limits_{1\leq n\leq b} \mu(E_n)\Big)^2}{\Big(\sum\limits_{1\leq n\leq b} \mu(E_n)\Big)^2 + (2\kappa+1) \sum\limits_{1\leq n\leq b} \mu(E_n)}\,.
\end{aligned}$$

By the divergent sum hypothesis, on letting $b \to \infty$ and $\lambda \to 0$, we obtain that

$$\limsup_{b\to\infty} \mu\Big(Z_{1,b} > \lambda \mathbb{E}(Z_{1,b})\Big) = 1\,,$$

which together with (23) completes the proof of the corollary. □

3. Establishing measure results for matrix transformations

3.1. Proof of Proposition 1

Let $X \subset \mathbb{R}^d$ be a compact set. The proof of Proposition 1 makes essential use of the work of Saussol [43] for general piecewise expanding maps $T : X \longrightarrow X$ on X. Clearly, our particular case in which $X = [0,1]^d$ and the real, non-singular matrix T (with the modulus of all eigenvalues strictly larger than one) that sends $\mathbf{x}$ to $T\mathbf{x}$ mod 1 defines a piecewise expanding map on X. In what follows will state and apply Saussol's results to our setup. So with this in mind, the first three parts, apart for the equivalence of the restricted measures $\mu|_{A_i}$ and $m_d|_{A_i}$ in part (ii), follow from [43, Theorem 5.1]. To prove the equivalence of the restricted measures we note that by [43, Proposition 5.1], the Randon-Nykodym derivative f of $\mu|_{A_i}$ with respect to $m_d|_{A_i}$ is m_d-almost surely strictly positive. Hence, for any measurable subset E, if $m_d|_{A_i}(E) > 0$, then

$$\mu|_{A_i}(E) = \int_E f(\mathbf{x}) m_d|_{A_i}(d\mathbf{x}) > 0$$

and so $m_d|_{A_i} \ll \mu|_{A_i}$. The other direction follows directly from part (i).

It remains to prove part (iv) of the proposition. Without loss of generality, we will assume that A_{ij} is A and thus $\mu|_{A_{ij}} = \mu$ and also μ is mixing with respect to T. The key to proving part (iv) is to use the fact that the acim μ satisfies the property of exponential decay of correlations. With this mind, for a set $E \subset \mathbb{R}^d$, define the *oscillation* of $\varphi \in L^1(m_d)$ over E as

$$\operatorname{osc}(\varphi, E) := \operatorname*{ess\text{-}sup}_E(\varphi) - \operatorname*{ess\text{-}inf}_E(\varphi).$$

For given real numbers $0 < \alpha \le 1$ and $0 < \epsilon_0 < 1$, define the following *α-seminorm*

$$|\varphi|_\alpha := \sup_{0<\epsilon\le\epsilon_0} \epsilon^{-\alpha} \int_{\mathbb{R}^d} \operatorname{osc}(\varphi, B(x,\epsilon))dx.$$

Let V_α be the space of $L^1(m_d)$-functions such that $|\varphi|_\alpha < \infty$ endowed with the norm

$$\|\varphi\|_\alpha := \|\varphi\|_{L^1(m_d)} + |\varphi|_\alpha.$$

Then $(V_\alpha, \|\cdot\|_\alpha)$ is a Banach space which does not depend on the choice of ϵ_0 and $V_\alpha \subset L^\infty(m_d)$ (see [43, Section 3]). Since the acim μ is mixing with respect to T, it follows from [43, Theorem 6.1] that there exist constants $C > 0$ and $0 \leq \gamma < 1$ such that for all $\psi \in V_\alpha$, for all $\phi \in L^1(\mu)$ and for all $n \in \mathbb{N}$, we have

$$\left| \int\limits_{[0,1]^d} \psi.\phi \circ T^n \, d\mu \; - \int\limits_{[0,1]^d} \psi \, d\mu \int\limits_{[0,1]^d} \phi \, d\mu \right| \;\leq\; C \;\|\psi\|_\alpha \;\|\phi\|_1 \;\gamma^n. \tag{24}$$

Now, let $\mathcal{C}$ be any collection of subsets of A satisfying the bounded property (**B**). Take $\psi = \chi_E$ and $\phi = \chi_F$ with $E, F \in \mathcal{C}$ and assume for the moment that

$$\sup_{E \in \mathcal{C}} \|\chi_E\|_\alpha < \infty. \tag{25}$$

Note that (25) implies that $\psi = \chi_E \in V_\alpha$ and thus together with (24), we obtain that

$$\left|\mu(E \cap T^{-n}F) - \mu(E)\mu(F)\right| \leq C \cdot \|\chi_E\|_\alpha \cdot \mu(F) \cdot \gamma^n \leq C \cdot \left(\sup_{E \in \mathcal{C}} \|\chi_E\|_\alpha \right) \cdot \gamma^n \cdot \mu(F).$$

Hence the exponentially mixing property (4) is satisfied for the collection $\mathcal{C}$. This proves part (iv) modulo (25).

We now prove (25). To start with, observe that for the characteristic function χ_E, the oscillation can only be non-zero on the boundary ∂E of E. It can be verified that for any $\epsilon > 0$,

$$\operatorname{osc}(\chi_E, B(x, \epsilon)) \;\leq\; \chi_{(\partial E)(\epsilon)}(x),$$

where $(\partial E)(\epsilon)$ is the ϵ-neighbourhood of ∂E. Thus,

$$\epsilon^{-\alpha} \int\limits_{\mathbb{R}^d} \operatorname{osc}(\chi_E, B(x, \epsilon)) dx \leq \epsilon^{-\alpha} \cdot m_d((\partial E)(\epsilon)).$$

By the bounded property (**B**) imposed on $\mathcal{C}$, there exists a constant C_1 such that the $(d-1)$-dimensional upper Minkowski content of ∂E

$$M^{*(d-1)}(\partial E) \leq C_1,$$

for all $E \in \mathcal{C}$. Hence, by the definition of $M^{*(d-1)}$, there exists a constant C_0 such that for all $E \in \mathcal{C}$

$$m_d((\partial E)(\epsilon)) \;<\; C_0 \epsilon.$$

Consequently, for all $E \in \mathcal{C}$,

$$|\chi_E|_\alpha = \sup_{0<\epsilon\le\epsilon_0} \epsilon^{-\alpha} \int_{\mathbb{R}^d} \operatorname{osc}(\chi_E, B(x,\epsilon))dx \le C_0\epsilon_0^{1-\alpha}.$$

On the other hand, for $E \subset [0,1]^d$, we have that

$$\|\chi_E\|_{L^1(m_d)} = m_d(E) \le 1.$$

Thus, for all $E \in \mathcal{C}$, it follows that

$$\|\chi_E\|_\alpha \le 1 + C_0\epsilon_0^{1-\alpha}.$$

Therefore, (25) holds and this completes the proof of the proposition. □

Remark 7. For the sake of completeness, we mention that in the case that T is an integer, non-singular, matrix transformation of the torus $\mathbb{T}^d$ with all eigenvalues in modulus strictly larger than one, Fan [15] proved the exponential decay of correlation formula (24). Also in the case $d = 1$, the first three parts of Proposition 1 coincide with the Main Result of Wagner [50]

3.2. Proof of Theorems 2, 3 and 5

Proof of Theorem 2. In view of parts (i) to (iii) of Proposition 1, if the acim μ has $\mathbb{T}^d$ as its support and is mixing with respect to T, then μ has one mixing component (namely the whole space $\mathbb{T}^d$) of period one. Furthermore, by part (iv) of Proposition 1, on this unique mixing component, μ is exponentially mixing with respect to $(T, \mathcal{C})$ for any collection $\mathcal{C}$ of subsets of $\mathbb{T}^d$ satisfying the bounded property (**B**). The desired counting part (9) of the theorem and the zero-full measure criteria (10) with respect to the measure μ now immediately follow on applying Theorem 1 and Corollary 1 respectively. To complete the proof of Theorem 2, it remains to prove that the measures μ and m_d share the same zero and full measure sets (note that we have already shown above that $\mu\big(W(T, \{E_n\})\big)$ is either 0 or 1). This follows directly from part (ii) of Proposition 1 since it implies that μ is equivalent to m_d and it is easily seen that equivalent measures share the same zero and full measure sets. □

The following lemma will be used in establishing Theorems 3 and 5.

Lemma 3. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that (i) all eigenvalues of T are of modulus strictly larger than $1 + \sqrt{d}$ or that (ii) T is integer and all eigenvalues of T are of modulus strictly larger than 1. Then there is a unique acim μ. Furthermore, such an acim μ has $\mathbb{T}^d$ as its support and is of maximal entropy.*

Proof. We are given that all eigenvalues of T are of modulus strictly larger than 1. Thus, by part (i) of Proposition 1, there exists an acim μ. By [9, Proposition 1], if all eigenvalues of T have modulus strictly larger than $1+\sqrt{d}$ or if T is integers, then the dynamical system $(\mathbb{T}^d, T)$ is topological transitive. Hence, by [9, Theorem 3 and its Corollary] the acim μ has the whole space $\mathbb{T}^d$ as its support and is the unique maximal entropy measure. $\square$

We now prove Theorems 3 and 5.

Proof of Theorem 3. By [9, Lemma 5], if the eigenvalues of T are all of modulus strictly larger than $1+\sqrt{d}$, then T is locally eventually onto. Thus the unique maximal entropy acim μ coming from Lemma 3 which has $\mathbb{T}^d$ as its support, is exact (see for example [38, Theorem 5.2.12]) and hence mixing with respect to T (see [37, Proposition 12.2]). In other words, this measure μ satisfies the hypotheses of Theorem 2. On applying Theorem 2, we obtain Theorem 3. $\square$

Proof of Theorem 5. Observe that if T is integer, then T is an endomorphism of the torus $\mathbb{T}^d$ (see for example, [51, Theorem 0.15]). By [51, Corollary 1.10.1 and Theorem 1.28], the Lebesgue measure m_d is mixing with respect to T. Furthermore, by [51, Theorem 8.15], m_d is of maximal entropy. Thus m_d is nothing but the unique maximal entropy acim μ in Lemma 3. Hence, Theorem 5 follows on applying Theorem 2 with $\mu = m_d$. $\square$

3.2.1. Proof of Corollary 3

As noted in Remark 5, the corollary follows directly from Theorem 5 on showing that the d-dimensional Lebesgue measure m_d of the hyperboloid region $H(\mathbf{a}, \psi(n))$ satisfies (14). It is easily versified that $m_d(H(\mathbf{a}, \psi(n)))$ is independent of the 'shift' $\mathbf{a} \in \mathbb{T}^d$. So with this in mind, it suffices to prove the following statement.

Lemma 4. *Given $d \in \mathbb{N}$ and $\delta > 0$, let*

$$H_d(\delta) := \left\{(x_1, \ldots, x_d) \in [0,1)^d : \|x_1\| \ldots \|x_d\| < \delta\right\}. \tag{26}$$

Then

$$m_d(H_d(\delta)) = \begin{cases} 1 & \text{if } \delta \geq 2^{-d} \\ 2^d\delta\left(\sum_{t=0}^{d-1} \frac{1}{t!}\left(\log \frac{1}{2^d\delta}\right)^t\right) & \text{if } \delta < 2^{-d}. \end{cases} \tag{27}$$

Proof. To simplify computations, first note that the measure of $H_d(\delta)$ is equal to 2^d times the measure of $[0, 1/2]^d \cap H_d(\delta)$. Furthermore, it is technically simpler to work with points restricted to $[0, 1/2]^d$ since the inequality under consideration is equivalent to

$$x_1 \ldots x_d \leq \delta. \tag{28}$$

So, from now on we will focus on computing the measure of the set

$$V_d(\delta) := \left\{(x_1, \ldots, x_d) \in [0,1/2]^d \,:\, x_1 \ldots x_d \le \delta\right\} \tag{29}$$

and recall that

$$2^d\, m_d\big(V_d(\delta)\big) \;=\; m_d\big(H_d(\delta)\big)\,. \tag{30}$$

Case (a): if $\delta \ge 2^{-d}$. Then it is a easily versified that $V_d(\delta) = [0,1/2]^d$ and this together with (30) implies (27). So, without loss of generality we can assume that $\delta < 2^{-d}$.

Case (b): if $\delta < 2^{-d}$. In view of (30), to establish (27) we need to show that for any $d \in \mathbb{N}$ and $0 < \delta < 2^{-d}$

$$m_d\big(V_d(\delta)\big) \;=\; \delta \sum_{t=0}^{d-1} \frac{1}{t!}\left(\log \frac{1}{2^d\delta}\right)^t . \tag{31}$$

This we now do by induction on d. For $d = 1$, we have that

$$m_1\big(V_1(\delta)\big) = m_1\big(\,\{x \in [0,1/2] \,:\, x \le \delta\}\,\big) \;=\; \delta \tag{32}$$

and this coincides with (31). Now let $d \ge 2$ and observe that we can rewrite (28) as

$$x_1 \ldots x_{d-1} \le \delta/x_d. \tag{33}$$

Note that since $(x_1, \ldots, x_d) \in [0,1/2]^d$, the left hand side of (33) is not bigger than $(1/2)^{d-1}$. Hence, it follows that for any $0 < x_d \le 2^{d-1}\delta$, inequality (33) is satisfied for all $0 \le x_1, \ldots, x_{d-1} \le 1/2$. The m_d-measure of the set of such points $(x_1, \ldots, x_d)$ is thus equal to $2^{-d+1} \times 2^{d-1}\delta = \delta$. On the other hand, for any fixed value of $x_d \in (2^{d-1}\delta, 1/2]$, the m_{d-1}-measure of the set of points $(x_1 \ldots x_{d-1}) \in [0,1/2]^{d-1}$ satisfying (33) is by definition equal to $m_{d-1}\big(V_{d-1}(\delta/x_d)\big)$. The upshot is that for any $d \ge 2$ and $0 < \delta < 2^{-d}$

$$m_d\big(V_d(\delta)\big) \;=\; \delta \,+ \int_{2^{d-1}\delta}^{1/2} m_{d-1}\big(V_{k-1}(\delta/x_d)\big)\, \mathrm{d}x_d. \tag{34}$$

Now assume that (31) holds with $d-1$ in place of d. Then, it follows via (34) that

$$\begin{aligned} m_d\big(V_d(\delta)\big) &= \delta \;+ \int_{2^{d-1}\delta}^{1/2} \frac{\delta}{x_d}\left(\sum_{t=0}^{d-2} \frac{1}{t!}\left(\log \frac{x_d}{2^{d-1}\delta}\right)^t\right) \mathrm{d}x_d \\ &= \delta \;+\; \delta \int_{1}^{\frac{1}{2^d\delta}} \left(\sum_{t=0}^{d-2} \frac{1}{t!}\frac{(\log y)^t}{y}\right) \mathrm{d}y \end{aligned}$$

$$
\begin{aligned}
&= \delta \;+\; \delta \sum_{t=0}^{d-2} \frac{1}{t!} \int_{1}^{\frac{1}{2^d\delta}} \frac{(\log y)^t}{y}\, \mathrm{d}y \\
&= \delta \;+\; \delta \sum_{t=0}^{d-2} \frac{1}{(t+1)!} \left(\log \frac{1}{2^d\delta}\right)^{t+1} \\
&= \delta \;+\; \delta \sum_{t=1}^{d-1} \frac{1}{t!} \left(\log \frac{1}{2^d\delta}\right)^{t} \;=\; \delta \sum_{t=0}^{d-1} \frac{1}{t!} \left(\log \frac{1}{2^d\delta}\right)^{t}.
\end{aligned}
$$

This completes the induction step and so establishes (31) for any $d \geq 1$. □

3.3. Proof of Theorem 4

We start by summarising various basic facts concerning β-transformations that will be utilized in proving Theorem 4. So, with this in mind, let β be a real number such that $|\beta| > 1$ and let $T_\beta : [0, 1) \to [0, 1)$ be the associated β-transformation given by

$$
T_\beta(x) = \beta x \pmod 1.
$$

For obvious reasons, when $\beta < -1$ the corresponding transformation is refereed to as the negative β-transformation.

For $\beta > 1$, Rényi [39, Theorem 1] proved that there exists a unique T_β-invariant measure μ_β (the so called Parry measure) that is strongly equivalent to (one-dimensional) Lebesgue measure m_1 on the unit interval. Clearly, this implies that μ_β is absolutely continuous with respect to Lebesgue measure. For the negative β-transformation, Ito and Sadahiro [26] proved that there is a unique T_β-invariant measure μ_β (the so called Yrrap measure) which is absolutely continuous with respect to Lebesgue measure m_1. The following proposition implies that μ_β is in fact strongly equivalent to m_1 when $\beta \leq -g$. Note that in view of [26, Theorem 16], for the negative β-transformation the corresponding density function of μ_β with respect to m_1 is given by

$$
h_\beta(x) := \frac{1}{F(\beta)} \sum_{\substack{n \geq 0 \\ T_\beta^n 1 \geq x}} \frac{1}{\beta^n},
$$

where

$$
F(\beta) := \int_0^1 \sum_{\substack{n \geq 0 \\ T_\beta^n 1 \geq x}} \frac{1}{\beta^n} dx
$$

is the normalising function.

Proposition 2. *Let $\beta \leq -g$. Then the Yrrap measure μ_β is strongly equivalent to the Lebesgue measure m_1 on the unit interval. More precisely, there exists a constant $C(\beta) > 0$ such that*

$$C(\beta)^{-1} \leq h_\beta(x) \leq C(\beta) \qquad \forall \quad x \in (0,1)\,.$$

Proof. For $\beta = -g$, it is easily verified that

$$h_\beta(x) = \begin{cases} \frac{1}{3+\beta} & \text{if} \quad 0 < x \leq 2+\beta \\ \\ \frac{-\beta}{3+\beta} & \text{if} \quad 2+\beta < x < 1. \end{cases}$$

Hence, we can choose $C(\beta) = \frac{1}{3+\beta}$. Without loss of generality, assume $\beta < -g$ and note that

$$1 + \sum_{n=0}^{\infty} \frac{1}{\beta^{2n+1}} \;=\; \frac{\beta^2+\beta-1}{\beta^2-1} \;\leq\; F(\beta) \;\leq\; \sum_{n=0}^{\infty} \frac{1}{\beta^{2n}} \;=\; \frac{\beta^2}{\beta^2-1}\,.$$

It then immediately follows that

$$h_\beta(x) \;\leq\; \frac{1}{F(\beta)} \sum_{n=0}^{\infty} \frac{1}{\beta^{2n}} \;=\; \frac{\beta^2}{\beta^2+\beta-1}$$

and

$$h_\beta(x) \;\geq\; \frac{1}{F(\beta)} \left(1 + \sum_{n=0}^{\infty} \frac{1}{\beta^{2n+1}} \right) \;=\; \frac{\beta^2+\beta-1}{\beta^2}.$$

Hence, we can choose $C(\beta) = \frac{\beta^2}{\beta^2+\beta-1}$. $\square$

The following result identifies the nature of the support of the T_β-invariant measure μ_β.

Proposition 3. *Let β be a real number with $|\beta| > 1$, μ_β be the associated Parry-Yrrap measure and let $K(\beta)$ denote the support of μ_β. Then*

$$K(\beta) = [0,1] \qquad \text{if} \quad \beta \in (-\infty, -g] \,\cup\, (1, +\infty)\,,$$

and $K(\beta)$ is a finite union of closed intervals contained in $[0,1]$ if $\beta \in (-g,-1)$. Furthermore, μ_β is mixing with respect to T_β and is equivalent to the measure $m|_{K(\beta)}$; i.e. the one-dimensional Lebesgue measure restricted to $K(\beta)$.

From this point onwards, given β with $|\beta| > 1$, μ_β will always denote the associated Parry-Yrrap measure and $K(\beta)$ will denote the support of μ_β.

Proof. If $\beta \in (-\infty, -g] \cup (1, +\infty)$, the result immediately follows from the fact that μ_β is strongly equivalent to the Lebesgue measure m_1 on $[0,1]$ – this is Proposition 2 for $\beta \leq -g$ and as already mentioned established by Rényi [39, Theorem 1] for $\beta > 1$. In general, Keller [28] proved for any $|\beta| > 1$ the support of μ_β is a finite union of closed intervals. The precise description of the closed intervals for the non-trivial case when $\beta \in (-g, -1)$ was given by Liao & Steiner [33, Theorem 2.1].

For the furthermore part, it follows via Rokhlin [40, Section 4.5] for $\beta > 1$ and Liao & Steiner [33, Corollary 2.3] for $\beta < -1$, that the T_β-invariant measure μ_β is exact and hence mixing with respect to T_β ([37, Proposition 12.2]). In turn, by the Main Result in [50] it follows that μ_β is equivalent to m_1 restricted to $K(\beta)$. $\square$

Of course, as indicated in above proof of the proposition, for $\beta \in (-\infty, -g] \cup (1, +\infty)$ we have that μ_β is strongly equivalent to the Lebesgue measure m_1 on $[0,1]$ rather than simply equivalent. The next statement is a straightforward consequence of Proposition 3 and basic properties of product measures.

Lemma 5. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal and all eigenvalues $\beta_1, \beta_2, \dots, \beta_d$ are of modulus strictly larger than 1. Then the product measure*

$$\nu := \mu_{\beta_1} \times \mu_{\beta_2} \times \cdots \times \mu_{\beta_d}$$

has support $K := \prod_{i=1}^d K(\beta_i)$ and is a T-invariant mixing measure that is equivalent to $m_d|_K$; i.e. the d-dimensional Lebesgue measure restricted to K.

Proof. By definition, the product measure ν has support $K = \prod_{i=1}^d K(\beta_i)$ and in view of Proposition 3 it is T-invariant and equivalent to the measure $m_d|_K$. Furthermore, since each μ_{β_i} is mixing with respect to T_{β_i}, on following the proof of [51, Theorem 1.24] it is easily verified that ν is mixing with respect to T. $\square$

We now show that Theorem 4 is an easy consequence of Lemma 5 together with Proposition 1 and Theorem 1. In the next section we will provide a self-contained and essentially elementary proof of Theorem 4 in the case the collection $\mathcal{C}$ of subsets of K is restricted to rectangles with sides parallel to the axes. This is an important class of "target sets" that clearly satisfy the bounded property (**B**) and the proof will avoid appealing to Proposition 1.

Proof of Theorem 4 using Proposition 1. Lemma 5 implies that product measure ν is a T-invariant mixing measure equivalent to $m_d|_K$. Hence by Proposition 1, ν is exponentially mixing with respect to $(T, \mathcal{C})$ where $\mathcal{C}$ is any collection of subsets E of K satisfying the bounded property (**B**). Then the main counting part of Theorem 4 immediately follows from Theorem 1. For the "furthermore" part we first recall (as we have done so several times) that by Rényi [39, Theorem 2] and Proposition 2, for any $\beta \in (-\infty, -g] \cup (1, +\infty)$ the Parry-Yrrap measure μ_β is strongly equivalent to the Lebesgue measure m_1 on $[0, 1]$. It thus follows that the product measure ν is strongly equivalent to m_d restricted to $K = \mathbb{T}^d$. The upshot of this is that we can replace ν by m_d in the first part of the theorem and thereby completes the proof of Theorem 4. $\square$

3.3.1. Theorem 4 for rectangles: a self contained and direct proof

In the proof of Theorem 4 given above, we make use of Proposition 1 to deduce that the product measure ν is exponentially mixing with respect to $(T, \mathcal{C})$ where $\mathcal{C}$ is any collection of subsets E of K satisfying the bounded property (**B**). The following result enables us to bypass the proposition in the case $\mathcal{C}$ is restricted to rectangles with sides parallel to the axes.

Lemma 6. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal and all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ are of modulus strictly larger than 1. Let $\nu := \mu_{\beta_1} \times \mu_{\beta_2} \times \cdots \times \mu_{\beta_d}$ be the product measure and $K := \prod_{i=1}^d K(\beta_i)$ be its support. Then ν is exponentially mixing with respect to $(T, \mathcal{R})$ for any collection $\mathcal{R}$ of rectangles of K with sides parallel to the axes.*

It is easily seen that by appealing to Lemma 6 instead of Proposition 1 in the "Proof of Theorem 4 using Proposition 1" given in the previous section, we obtain the special case of Theorem 4 in which $\mathcal{C}$ is any collection $\mathcal{R}$ of rectangles of K with sides parallel to the axes. In other words, it enables us to provide a self-contained and direct proof of Theorem 4 for rectangular target sets.

Proof of Lemma 6. First, we assert that for any $\beta \in \mathbb{R}$ with $|\beta| > 1$, μ_β is exponentially mixing with respect to $(T_\beta, \mathcal{C})$ where $\mathcal{C}$ is any collection of intervals of $K(\beta)$. When $\beta > 1$, this is the classical result of Gel'fond [21, Formula (12)] and Philipp [42, Lemma 7]. When $\beta < -1$, the assertion follows from a general result of Baladi [1, Theorem 3.4] for piecewise monotone expanding interval maps.

We now verify that the product ν satisfies the desired exponentially mixing property. With this in mind, let $E = B(z_1, r_1) \times \cdots \times B(z_d, r_d)$ and $F = B(z_1', r_1') \times \cdots \times B(z_d', r_d')$ be any two rectangles in $\mathcal{R}$. Then

$$\nu(E \cap T^{-n}F) = \int \chi_E(x_1, \ldots, x_d)\, \chi_F(T_{\beta_1}^n x_1, \ldots, T_{\beta_d}^n x_d)\; d\mu_{\beta_1}(x_1) \cdots\, d\mu_{\beta_d}(x_d)\,,$$

and by the property of the product measure the right hand side equals

$$\int \chi_{B(z_1,r_1)}(x_1)\chi_{B(z_1',r_1')}(T_{\beta_1}^n x_1) d\mu_{\beta_1}(x_1) \cdots \int \chi_{B(z_d,r_d)}(x_d)\chi_{B(z_d',r_d')}(T_{\beta_d}^n x_d) d\mu_{\beta_d}(x_d).$$

It then follows by the exponentially mixing property of T_{β_i}, that

$$\nu(E \cap T^{-n}F) = \prod_{i=1}^{d} \Big(\mu_{\beta_i}(B(z_i, r_i))\, \mu_{\beta_i}(B(z_i', r_i')) + O(\gamma_i^n)\, \mu_{\beta_i}(B(z_i', r_i')) \Big)$$

where $0 < \gamma_i < 1$. This together with the fact that

$$\nu(E) = \mu_{\beta_1}(B(z_1, r)) \cdots \mu_{\beta_d}(B(z_d, r)), \quad \nu(F) = \mu_{\beta_1}(B(z_1', r')) \cdots \mu_{\beta_d}(B(z_d', r'))$$

and $\mu_{\beta_i}(B(z_i, r)) \le 1$ $(1 \le i \le d)$ implies that

$$\nu(E \cap T^{-n}F) = \nu(E)\nu(F) + O(\gamma^n)\mu(F) \quad \text{where} \quad \gamma = \max\{\gamma_1, \ldots, \gamma_d\}.$$

In other words, (4) holds for rectangles in $\mathcal{R}$ and we are done. $\square$

3.3.2. Extending Corollary 2 to incorporate eigenvalues in $[-1, 1]$

We show that on assuming $\psi(n) \to 0$ as $n \to \infty$, we can naturally extend Corollary 2 to the situation that all the eigenvalues of T are in $(-\infty, -g] \cup [-1, +\infty)$; that is to say, we can incorporate the interval $[-1, 1]$. In short, assuming that T has eigenvalues in $[-1, 1]$, we do this in most cases by reformulating the shrinking target set $W(T, \psi, \mathbf{a})$ in terms of related "lower dimensional' sets $W(T_*, \psi, \mathbf{a}_*)$ for which Corollary 2 in its current form is applicable. In other words, all the eigenvalues of the related transformation T_* are in $(-\infty, -g] \cup (1, +\infty)$. Let T be a real, non-singular, diagonal matrix transformation of the torus $\mathbb{T}^d$ with eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ in $(-\infty, -g] \cup [-1, +\infty)$. Without loss of generality, assume that there is at least one eigenvalue in $[-1, 1]$. In fact, let us assume that there is only one such eigenvalue, say β_1. It should be self-evident how to deal with the situation in which that are multiple eigenvalues in $[-1, 1]$. We consider the three separate situations depending on whether $|\beta_1| < 1$, $\beta_1 = 1$ or $\beta_1 = -1$. Note that $T = \mathrm{diag}(\beta_1, \cdots, \beta_d)$ and so for any $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{T}^d$

$$T^n(\mathbf{x}) = (T^n_{\beta_1}(x_1), T^n_{\beta_2}(x_2), \cdots, T^n_{\beta_d}(x_d)).$$

(i) *We assume* $|\beta_1| < 1$. We distinguish between three subcases:

- Case 1: $\beta_1 = 0$. Then it is easily verified that

$$W(T, \psi, \mathbf{a}) = \begin{cases} \emptyset & \text{if } a_1 \neq 0 \\ \mathbb{T} \times W(T_*, \psi, \mathbf{a}_*) & \text{if } a_1 = 0, \end{cases}$$

  where

$$T_* := \mathrm{diag}(\beta_2, \cdots, \beta_d) \quad \text{and} \quad \mathbf{a}_* := (a_2, \cdots, a_d) \in \mathbb{T}^{d-1}. \tag{35}$$

- Case 2: $0 < \beta_1 < 1$. Then $T^n_{\beta_1}(x_1) = \beta_1^n x_1 \to 0$ as $n \to \infty$ for any $x_1 \in \mathbb{T}$. Note that zero is the unique fixed point of T_{β_1}. Thus,

$$\{x_1 \in [0,1) : T^n_{\beta_1}(x_1) < \psi(n) \ \text{ for infinitely many } n \in \mathbb{N}\} \ = \ I_* \ \text{ or } \ \overline{I_*},$$

where $I_* := [0, \min\{1, \tau\})$, $\overline{I_*}$ is the closure of the set I_* and

$$\tau := \limsup_{n\to\infty} \psi(n)\,|\beta_1|^{-n}. \tag{36}$$

Indeed, the set under consideration is I_* if $\psi(n)\,|\beta_1|^{-n} \le \tau$ for infinitely many $n \in \mathbb{N}$ and $\overline{I_*}$ otherwise. Hence, with T_* and $\mathbf{a}_*$ as in (35), it follows that

$$W(T,\psi,\mathbf{a}) = \begin{cases} \emptyset & \text{if } \ a_1 \neq 0 \\ \{0\} \times W(T_*,\psi,\mathbf{a}_*) & \text{if } \ a_1 = 0 \text{ and } \tau = 0, \\ I_* \times W(T_*,\psi,\mathbf{a}_*) \quad \text{or} \quad \overline{I_*} \times W(T_*,\psi,\mathbf{a}_*) & \text{if } \ a_1 = 0 \text{ and } \tau > 0. \end{cases}$$

- Case 3: $-1 < \beta_1 < 0$. Let F be the countable set of all preimages of zero; that is,

$$F := \{x_1 \in \mathbb{T} : T^n_{\beta_1}(x_1) = 0 \text{ for some } n \ge 0\}.$$

If $x_1 \notin F$, then $T_{\beta_1} x_1 = \beta_1 x_1 + 1$ and

$$T^n_{\beta_1} x_1 = \beta_1^n x_1 + \beta_1^{n-1} + \beta_1^{n-2} + \cdots + \beta_1 + 1 \to \frac{1}{1-\beta_1}$$

as $n \to \infty$. Note that zero and $\frac{1}{1-\beta_1}$ are the fixed points of T_{β_1}. Thus,

$$\{x_1 \in [0,1) \setminus F : T^n_{\beta_1}(x_1) < \psi(n) \ \text{ for infinitely many } n \in \mathbb{N}\} = J_* \setminus F \ \text{ or } \ \overline{J_*} \setminus F,$$

where $J_* := \Big((\frac{1}{1-\beta_1} - \tau, \frac{1}{1-\beta_1} + \tau) \cap [0,1) \Big)$, $\overline{J_*}$ is the closure of the set J_* and τ is given by (36). Indeed, the set under consideration is $J_* \setminus F$ if $\psi(n)\,|\beta_1|^{-n} \le \tau$ for infinitely many $n \in \mathbb{N}$ and $\overline{J_*} \setminus F$ otherwise. Hence, with T_* and $\mathbf{a}_*$ as in (35), it follows that

$$W(T,\psi,\mathbf{a}) = \begin{cases} F \times W(T_*,\psi,\mathbf{a}_*) & \text{if } \ a_1 = 0 \\ \emptyset & \text{if } \ a_1 \neq \frac{1}{1-\beta_1}, a_1 \neq 0 \\ \{\frac{1}{1-\beta_1}\} \times W(T_*,\psi,\mathbf{a}_*) & \text{if } \ a_1 = \frac{1}{1-\beta_1} \text{ and } \tau = 0, \\ J_* \times W(T_*,\psi,\mathbf{a}_*) \quad \text{or} \quad \overline{J_*} \times W(T_*,\psi,\mathbf{a}_*) & \text{if } \ a_1 = \frac{1}{1-\beta_1} \text{ and } \tau > 0. \end{cases}$$

The upshot is that in order to determine the size of $W(T,\psi,\mathbf{a})$ or the behaviour of the associated counting function we need to investigate the shrinking target set $W(T_*,\psi,\mathbf{a}_*) \subseteq \mathbb{T}^{d-1}$. Recall, that for ease of discussion we are assuming that β_1 is the only eigenvalue of T in $[-1,1]$. Thus, all the eigenvalues of the related transformation T_* are in $(-\infty,-g] \cup (1,+\infty)$ and so Corollary 2 is applicable with T replaced by T_*, $\mathbf{a}$ replaced by $\mathbf{a}_*$ and d replaced by $d-1$.

(ii) *We assume* $\beta_1 = 1$. For any $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{T}^d$ and $\mathbf{a} = (a_1, \cdots, a_d) \in \mathbb{T}^d$, the condition $T^n(\mathbf{x}) \in B(\mathbf{a}, \psi(n))$ implies that

$$\|x_1 - a_1\| < \psi(n).$$

Now, since we are assuming that $\psi(n) \to 0$ and $n \to \infty$, it follows that for any $\mathbf{x} \in W(T, \psi, \mathbf{a})$ we must have that $x_1 = a_1$. Hence, with T_* and $\mathbf{a}_*$ as in (35), it follows that

$$W(T, \psi, \mathbf{a}) = \{a_1\} \times W(T_*, \psi, \mathbf{a}_*)\,.$$

As in (i), the upshot is that we need to investigate the shrinking target set $W(T_*, \psi, \mathbf{a}_*) \subseteq \mathbb{T}^{d-1}$ and that for this setup Corollary 2 is applicable with T replaced by T_*, $\mathbf{a}$ replaced by $\mathbf{a}_*$ and d replaced by $d-1$.

(iii) *We assume* $\beta_1 = -1$. Then, for $x_1 \neq 0$

$$T^n_{\beta_1}(x_1) = \begin{cases} x_1 & \text{if } \ n \text{ is even} \\ -x_1 + 1 & \text{if } \ n \text{ is odd.} \end{cases}$$

If $a_1 = 0$, then the same reasoning as in (ii) shows that $W(T, \psi, \mathbf{a}) = \{0\} \times W(T_*, \psi, \mathbf{a}_*)$ and so we can apply Corollary 2 with T replaced by T_*, $\mathbf{a}$ replaced by $\mathbf{a}_*$ and d replaced by $d-1$. If $a_1 \neq 0$, it follows that for any $\mathbf{x} \in W(T, \psi, \mathbf{a})$ we must have that $x_1 = a_1$ or $x_1 = -a_1 + 1$. Thus, with T_* and $\mathbf{a}_*$ as in (35), we have that

$$W(T, \psi, \mathbf{a}) = \{a_1\} \times W'(T_*, \psi, \mathbf{a}_*) \bigcup \{1 - a_1\} \times W''(T_*, \psi, \mathbf{a}_*),$$

where

$$\begin{aligned} W'(T_*, \psi, \mathbf{a}_*) &:= \{\mathbf{x} \in \mathbb{T}^{d-1} : T_*^n(\mathbf{x}) \in B(\mathbf{a}_*, \psi(n)) \ \text{ for infinitely many even } n \in \mathbb{N}\}, \\ W''(T_*, \psi, \mathbf{a}_*) &:= \{\mathbf{x} \in \mathbb{T}^{d-1} : T_*^n(\mathbf{x}) \in B(\mathbf{a}_*, \psi(n)) \ \text{ for infinitely many odd } n \in \mathbb{N}\}. \end{aligned}$$

Now observe that $W'(T_*, \psi, \mathbf{a}_*)$ and $W''(T_*, \psi, \mathbf{a}_*)$ are shrinking target sets with respect to the transformation $T_* \circ T_*$ of the torus $\mathbb{T}^{d-1}$. Indeed,

$$W'(T_*, \psi, \mathbf{a}_*) = \{\mathbf{x} \in \mathbb{T}^{d-1} : (T_* \circ T_*)^n(x) \in B(\mathbf{a}_*, \psi(2n)) \ \text{ for infinitely many } n \in \mathbb{N}\},$$

and

$$\begin{aligned} W''(T_*, \psi, \mathbf{a}_*) = &\{\mathbf{x} \in \mathbb{T}^{d-1} : (T_* \circ T_*)^n(x) \in T_*^{-1} B(\mathbf{a}_*, \psi(2n+1)) \\ &\text{for infinitely many } n \in \mathbb{N}\}. \end{aligned}$$

Now in essence, the above procedure removes the presence of the problematic eigenvalue $\beta_1 = -1$ but still none of the results we have established for matrix transformations are applicable to the above shrinking targets sets. The reason for this is simple. The composition map $T_* \circ T_*$ is not a matrix transformation of the torus $\mathbb{T}^{d-1}$. The upshot is that we need to appeal to Theorem 1 and its corollary directly. In view of the argument set out in §3.3.1 that provides a self-contained proof of Theorem 4 for rectangular target sets, it is easily seen that the desired counting and measure statements for $W'(T_*, \psi, \mathbf{a}_*)$ and $W''(T_*, \psi, \mathbf{a}_*)$ would follow on showing that the product measure ν is exponentially mixing with respect to $(T_* \circ T_*, \mathcal{R})$. To establish the latter, we first recall (see the start of the proof of Lemma 6) that for any $\beta \in \mathbb{R}$ with $|\beta| > 1$, μ_β is exponentially mixing with respect to $(T_\beta, \mathcal{C})$ where $\mathcal{C}$ is any collection of intervals of $K(\beta)$. Hence by definition, there exists a constant $0 < \gamma < 1$ such that

$$\mu_\beta\big(E \cap T^{-2n} F\big) = \mu_\beta(E)\mu_\beta(F) + O(\gamma^{2n})\, \mu_\beta(F)$$

for any $E, F \in \mathcal{C}$. In other words,

$$\mu_\beta\big(E \cap (T \circ T)^{-n} F\big) = \mu_\beta(E)\mu_\beta(F) + O\big((\gamma^2)^n\big)\, \mu_\beta(F),$$

and so μ_β is exponentially mixing with respect to $(T_\beta \circ T_\beta, \mathcal{C})$. Then, on mimicking the proof of Lemma 6 with T replaced by $T_* \circ T_*$ and d replaced by $d-1$, we conclude that ν is exponentially mixing with respect to $(T_* \circ T_*, \mathcal{R})$ for any collection $\mathcal{R}$ of rectangles of K with sides parallel to the axes.

3.3.3. A nifty "reduction" argument

In this section we show that when T is an integer matrix transformation, the diagonal assumption in Theorem 4 can be relaxed to T is diagonalizable over $\mathbb{Z}$. So, suppose T is diagonalizable over $\mathbb{Z}$. Then by definition, there exist a nonsingular integer matrix P and a diagonal integer matrix D such that product matrix relationship $P \cdot T = D \cdot P$ holds. In turn, there exists an invertible mapping

$$\phi : \mathbb{T}^d \to \mathbb{T}^d \quad \text{ such that } \quad \phi \circ T = D \circ \phi\,. \tag{37}$$

Obviously, the diagonal entries of D are the eigenvalues of T and we assume these integers are of modulus strictly larger than 1.

Now recall that for diagonal transformations such as D, the "Proof of Theorem 4 using Proposition 1" makes key use of the fact that the product measure ν is D-invariant and is exponentially mixing with respect to $(D, \mathcal{C})$ where $\mathcal{C}$ is any collection of subsets E of K satisfying the bounded property (**B**). We claim that the image measure $\nu \circ \phi$ is T-invariant and exponentially mixing with respect to $(D, \mathcal{C})$. The proof of Theorem 4 can then be modified in the obvious manner to deal with the more general (integer) situation in which T is diagonalizable over $\mathbb{Z}$. To establish the claim, first note that for any measurable set $A \subset \mathbb{T}^d$, on using the fact that ν is D-invariant it follows that

$$\nu \circ \phi(T^{-1}A) = \nu(\phi(T^{-1}A)) = \nu(D^{-1}(\phi(A))) = \nu(\phi(A)) = \nu \circ \phi(A).$$

Thus, $\nu \circ \phi$ is T-invariant. Next, since ϕ is linear the (inverse) image of any collection $\mathcal{C}$ of subsets E of K satisfying the bounded property (**B**) also satisfies the bounded property (**B**). Hence, for any $E, F \in \mathcal{C}$, noting that $\phi(E \cap T^{-n}F) = \phi(E) \cap \phi(T^{-n}F)$, it follows that there exists a constant $0 < \gamma < 1$ such that

$$\begin{aligned}\nu \circ \phi(E \cap T^{-n}F) &= \nu\big(\phi(E \cap T^{-n}F)\big) = \nu\big(\phi(E) \cap \phi(T^{-n}F)\big) \\ &= \nu\big(\phi(E) \cap D^{-n} \circ \phi(F)\big) \\ &= \nu\big(\phi(E)\big)\, \nu\big(\phi(F)\big) \;+\; O\big(\gamma^n\big)\; \nu\big(\phi(F)\big) \\ &= \nu \circ \phi(E)\; \nu \circ \phi(F) \;+\; O(\gamma^n)\; \nu \circ \phi(F)\end{aligned}$$

as desired. Note that the third displayed line uses the fact that ν is exponentially mixing with respect to $(D, \mathcal{C})$.

Remark 8. We remark that for real non-integer matrices, we cannot in general use the above argument to extend Theorem 4 to the situation that T is diagonalizable over $\mathbb{Z}$. In short, the commutative property $\phi \circ T = D \circ \phi$ may not be true since the product matrix relationship $P \cdot T = D \cdot P$ does not guarantee that

$$P(T\mathbf{x} \bmod 1) \bmod 1 = T(P\mathbf{x} \bmod 1) \bmod 1.$$

4. Establishing dimension results for matrix transformations

We begin with a brief account in which we bring together various statements concerning Hausdorff measure and dimension that we will utilise in the course of establishing Theorems 6 and 7.

4.1. Preliminaries

We start by defining Hausdorff measure and dimension for completeness and for establishing some notation. Let X be a subset of $\mathbb{R}^d$. For $\rho > 0$, a countable collection $\{B_i\}$ of Euclidean balls in $\mathbb{R}^d$ of diameter $d_i \leq \rho$ for each i such that $X \subset \bigcup_i B_i$ is called a ρ-cover for X. Let s be a non-negative number and define

$$\mathcal{H}^s_\rho(X) \;=\; \inf \left\{ \sum_i d_i^s \;\; : \{B_i\} \text{ is a } \rho\text{−cover of } X \right\} ,$$

where the infimum is taken over all possible ρ-covers of X. The *s-dimensional Hausdorff measure* $\mathcal{H}^s(X)$ of X is defined by

$$\mathcal{H}^s(X) = \lim_{\rho \to 0} \mathcal{H}^s_\rho(X) = \sup_{\rho > 0} \mathcal{H}^s_\rho(X)$$

and the *Hausdorff dimension* $\dim X$ of X by

$$\dim_{\mathrm{H}} X = \inf\{s : \mathcal{H}^s(X) = 0\} = \sup\{s : \mathcal{H}^s(X) = \infty\} .$$

Further details and alternative definitions of Hausdorff measure and dimension can be found in [14,35]. It is easily verified (see [14, Corollary 2.4]) that the Hausdorff dimension of a set is invariant under bi-Lipschitz maps.

Lemma 7. *Let X be a subset of $\mathbb{R}^d$ and $f : X \to \mathbb{R}^d$ be a bi-Lipschitz map; i.e.*

$$c_1 |x - y| \leq |f(x) - f(y)| \leq c_2 |x - y| \qquad (x, y \in X)$$

where $0 < c_1 \leq c_2 < \infty$, then $\dim_{\mathrm{H}} f(X) = \dim_{\mathrm{H}} X$.

We now describe a deep and powerful mechanism for obtaining lower bounds for the Hausdorff dimension of a large class of "rectangular" lim sup sets.

4.1.1. Mass transference principle for rectangles

The discussion below is tailored to the application we have in mind. It is far from the most general and powerful setup of the Mass Transference Principle. We begin by describing the original 'balls to balls' principle which is all that is required for directly proving Theorem 6 and Theorem 7. However, we will deduce Theorem 6 from a more general statement concerning rectangular target sets and for this we will require the more versatile 'rectangle to rectangle' principle.

To set the scene, let X be a locally compact subset of $\mathbb{R}^d$ equipped with a non-atomic probability measure μ. Suppose there exist constants $\delta > 0$, $0 < a \leq 1 \leq b < \infty$ and $r_0 > 0$ such that

$$a\,r^{\delta} \;\leq\; \mu(B) \;\leq\; b\,r^{\delta} \tag{38}$$

for any ball $B = B(x, r)$ with $x \in X$ and radius $r \leq r_0$. Such a measure is said to be *δ-Ahlfors regular*. It is well known that if X supports a δ-Ahlfors regular measure μ, then $\dim_{\mathrm{H}} X = \delta$ and moreover that μ is strongly equivalent to δ-dimensional Hausdorff measure $\mathcal{H}^{\delta}$ – see [14,35] for the details. The latter implies that (38) is valid with μ replaced by $\mathcal{H}^s$. Next, given $s > 0$ and a ball $B = B(x, r)$ we define the scaled ball

$$B^s := B\big(x, r^{\frac{s}{\delta}}\big) ,$$

and so by definition $B^{\delta} = B$. The following Mass Transference Principle [3] allows us to transfer $\mathcal{H}^{\delta}$-measure theoretic statements for lim sup subsets of X to general $\mathcal{H}^s$-measure theoretic statements.

Theorem 10 *(MTP: balls to balls). Let X be a locally compact subset of $\mathbb{R}^d$ equipped with a δ-Ahlfors regular measure μ. Let $\{B_n\}_{n\in\mathbb{N}}$ be a sequence of balls in X with radius $r(B_n) \to 0$ as $n \to \infty$. Let $s \geq 0$ and suppose that*

$$\mathcal{H}^{\delta}\big(\limsup_{n\to\infty} B_n^s\big) = \mathcal{H}^{\delta}(X).$$

Then,

$$\mathcal{H}^{s}\big(\limsup_{n\to\infty} B_n\big) = \mathcal{H}^{s}(X).$$

Note that by the definition of Hausdorff dimension, Theorem 10 implies that

$$\dim_{\mathrm{H}}\big(\limsup_{n\to\infty} B_n\big) \geq s\,, \tag{39}$$

and moreover that $\mathcal{H}^s(\limsup_{n\to\infty} B_n) = \infty$ if $s < \delta$. We now describe a recent result due to Wang & Wu [52] that gives a lower bound for the Hausdorff dimension of lim sup sets defined via rectangles rather than just balls. So with this in mind, fix an integer $p \geq 1$ and for $1 \leq i \leq p$, let X_i be a subset of $\mathbb{R}^{d_i}$. Obviously, if B_i is a ball in X_i then $\prod_{i=1}^p B_i$ is in general a rectangle in the product space $\prod_{i=1}^p X_i$. The following statement for lim sup sets arising from sequences of such rectangles is a much simplified version of [52, Theorem 3.4]. As we shall see, it is more than adequate for our purpose.

Theorem 11 *(MTP: rectangles to rectangles). For each $1 \leq i \leq p$, let X_i be a locally compact subset of $\mathbb{R}^{d_i}$ equipped with a δ_i-Ahlfors regular measure μ_i. Let $\{B_{i,n}\}_{n\in\mathbb{N}}$ be a sequence of balls in X_i with radius $r(B_{i,n}) \to 0$ as $n \to \infty$ for each $1 \leq i \leq p$ and assume that there exist $\mathbf{v} = (v_1, \dots, v_p) \in (\mathbb{R}^+)^p$ and a sequence $\{r_n\}_{n\in\mathbb{N}}$ of positive real numbers such that*

$$r(B_{i,n}) = r_n^{v_i} \quad \text{for all } 1 \leq i \leq p. \tag{40}$$

Suppose that there exists $(s_1, \dots, s_p) \in \prod_{i=1}^p (0, \delta_i)$ such that

$$\mu_1 \times \cdots \times \mu_p\Big(\limsup_{n\to\infty} \textstyle\prod_{i=1}^p B_{i,n}^{s_i}\Big) = \mu_1 \times \cdots \times \mu_p\Big(\textstyle\prod_{i=1}^p X_i\Big). \tag{41}$$

Then, we have that

$$\dim_{\mathrm{H}}\big(\limsup_{n\to\infty} \textstyle\prod_{i=1}^p B_{i,n}\big) \geq \displaystyle\min_{1\leq i\leq p} s(\mathbf{u}, \mathbf{v}, i),$$

where $\mathbf{u} = (u_1, \dots, u_p)$ with $u_i = s_i v_i/\delta_i$ for $1 \leq i \leq p$, and

$$s(\mathbf{u}, \mathbf{v}, i) := \sum_{k\in\mathcal{K}_1(i)} \delta_k + \sum_{k\in\mathcal{K}_2(i)} \delta_k\big(1 - \frac{v_k - u_k}{v_i}\big) + \sum_{k\in\mathcal{K}_3(i)} \frac{u_k\delta_k}{v_i}\,,$$

with the sets

$$\mathcal{K}_1(i) := \{1 \leq k \leq p : u_k \geq v_i\}, \qquad \mathcal{K}_2(i) := \{1 \leq k \leq p : v_k \leq v_i\},$$

and

$$\mathcal{K}_3(i) := \{1, \ldots, p\} \setminus (\mathcal{K}_1(i) \cup \mathcal{K}_2(i))$$

forming a partition of $\{1, \ldots, p\}$.

Note that since the radius $r(B_{i,n}) \to 0$ as $n \to \infty$ for each $1 \le i \le p$ we automatically have that $\lim_{n\to\infty} r_n = 0$. Also note that if (40) holds for some $\mathbf{v} = (v_1, \ldots, v_p) \in (\mathbb{R}^+)^p$, then it holds for $c\mathbf{v} = (cv_1, \ldots, cv_p)$ where $c > 0$ is a constant. Thus, the choice of $\mathbf{v}$ and therefore $\mathbf{u}$ is not unique. However, it is easily seen that the "dimension number" $s(\mathbf{u}, \mathbf{v}, i)$ is not effected; i.e.

$$s(\mathbf{u}, \mathbf{v}, i) = s(c\mathbf{u}, c\mathbf{v}, i)\,.$$

For the sake of convenience, we refer to a collection of rectangles $\{R_n := \prod_{i=1}^p B_{i,n}\}_{n\in\mathbb{N}}$ with sidelengths satisfying (40) as a collection of *rectangles with exponent* $\mathbf{v}$. Thus, Theorem 11 can be regarded as the dimension analogue of the original Mass Transference Principle for lim sup sets arising from rectangles with exponent $\mathbf{v}$.

Remark 9. It is worth mentioning that assumption (40) can be easily weakened to the following statement: for any $1 \le i, j \le p$,

$$\lim_{n\to\infty} \frac{\log r(B_{i,n})}{\log r(B_{j,n})} \qquad \text{exists and is finite.}$$

To see this, let $r_n = r(B_{1,n})$. Then, there exists $\mathbf{v} = (v_1, \ldots, v_p) \in (\mathbb{R}^+)^p$ and sequences $\{v_{i,n}\}_{n\in\mathbb{N}}$ for each $1 \le i \le p$, such that

$$r(B_{i,n}) = r_n^{v_{i,n}} \quad \text{with} \quad \lim_{n\to\infty} v_{i,n} = v_i\ .$$

Now, given $\epsilon > 0$ consider the associated lim sup set R^*_ϵ obtained by replacing the balls $B_{i,n}$ by balls $B^*_{i,n}$ with the same centre but radius $r(B^*_{i,n}) = r_n^{v_i+\epsilon}$ $(1 \le i \le p)$. Then,

$$R^*_\epsilon := \limsup_{n\to\infty} \textstyle\prod_{i=1}^p B^*_{i,n} \subset \limsup_{n\to\infty} \prod_{i=1}^p B_{i,n}$$

and on applying Theorem 11 we obtain a lower bound for $\dim_{\mathrm{H}} R^*_\epsilon$ which converges to the desired dimensional number as $\epsilon \to 0$.

Remark 10. In the above setup the rectangles arise as products of balls B_i in X_i. Balls of radius ρ are of course ρ-neighbourhoods of special points; namely their centres. The general form of the Mass Transference Principle of Wang & Wu is based on the framework of ubiquitous systems. This allows them to naturally consider the situation in which the balls B_i are replaced by ρ-neighbourhoods of special sets called resonant sets. It is easily verified that the κ-scaling property for resonant sets within the general dimension statement [52, Theorem 3.4] is satisfied with $\kappa = 0$ when the resonant sets are points. In turn, this together with [52, Proposition 3.1] directly yields Theorem 11. It is also worth mentioning that if we replace the full measure condition (41) by the stronger 'local ubiquity system for rectangles' condition [52, Definition 3.2] and also assume that the radii of the balls in the given sequence are non-increasing, then we are able to conclude [52, Theorem 3.2] that

$$\mathcal{H}^s\big(\limsup_{n\to\infty}\textstyle\prod_{i=1}^p B_{i,n}\big) = \mathcal{H}^s\big(\textstyle\prod_{i=1}^p X_i\big) \quad \text{with} \quad s = \displaystyle\min_{1\le i\le p} s(\mathbf{u},\mathbf{v},i).$$

In other words, we obtain a complete 'rectangle to rectangle' analogue of the original Mass Transference Principle.

It is easily seen that in the case of balls, Theorem 11 coincides with the dimension statement (39). Indeed, when $p = 1$ we have that that $\mathcal{K}_1(1) = \mathcal{K}_3(1) = \emptyset$ and $\mathcal{K}_2(1) = \{1\}$. Hence, it follows that

$$s(\mathbf{u},\mathbf{v},1) = \delta_1\big(1 - \frac{v_1 - u_1}{v_1}\big) = \frac{\delta_1 u_1}{v_1} = s_1\,,$$

and so Theorem 11 implies that

$$\dim_{\mathrm{H}} \limsup_{n\to\infty} B_{1,n} \ge s_1$$

as claimed.

4.2. Proof of Theorem 6 via a dimension theorem for rectangular targets

As mentioned in Remark 6, straight after the statement of Theorem 6, we deduce the theorem from a more general statement concerning rectangular target sets. This we now describe and prove. For $1 \le i \le d$, let $\psi_i : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive function. For convenience, let $\Psi := (\psi_1, \ldots, \psi_d)$ and for $n \in \mathbb{N}$ let $\Psi(n) := (\psi_1(n), \ldots, \psi_d(n))$. Fix some point $\mathbf{a} := (a_1, \ldots, a_d) \in \mathbb{T}^d$ and for $n \in \mathbb{N}$, let

$$R\big(\mathbf{a}, \Psi(n)\big) := \Big\{\mathbf{x} \in \mathbb{T}^d : \|x_i - a_i\| \le \psi_i(n)\ (1 \le i \le d)\Big\}.$$

Clearly, $R\big(\mathbf{a}, \Psi(n)\big)$ is a rectangle centred at the fixed point $\mathbf{a}$. In turn, let

$$W(T, \Psi, \mathbf{a}) := \big\{\mathbf{x} \in \mathbb{T}^d : T^n(\mathbf{x}) \in R\big(\mathbf{a}, \Psi(n)\big) \ \text{ for infinitely many } n \in \mathbb{N}\big\}\,.$$

It is evident that the family of rectangular target sets $\{R\big(\mathbf{a}, \Psi(n)\big)\}_{n\ge1}$ satisfy the bounded property ($\mathbf{B}$) and indeed the stronger property ($\mathbf{P_a}$) based on Gallagher's property (**P**) as described in §5.1. Note that when $\psi := \psi_1 = \cdots = \psi_d$, the rectangles are squares and so

$$W(T, \Psi, \mathbf{a}) = W(T, \psi, \mathbf{a})\,.$$

Also, it is easily verified that if

$$T = \mathrm{diag}\,(\beta_1, \ldots, \beta_d) \qquad (\beta_i \in \mathbb{R})$$

then

$$W(T, \Psi, \mathbf{a}) = \left\{\mathbf{x} \in \mathbb{T}^d : |T^n_{\beta_i} x_i - a_i| \leq \psi_i(n)\ (1 \leq i \leq d) \ \text{ for infinitely many } n \in \mathbb{N}\right\}, \tag{42}$$

where T_{β_i} is the standard β-transformation with $\beta = \beta_i$.

It turns out that the Hausdorff dimension of the shrinking target set $W(T, \Psi, \mathbf{a})$ is dependent on the set $\mathcal{U}(\Psi)$ of accumulation points $\mathbf{t} = (t_1, t_2, \ldots, t_d)$ of the sequence $\big\{(-\frac{\log\psi_1(n)}{n}, \cdots, -\frac{\log\psi_d(n)}{n})\big\}_{n\geq 1}$.

Theorem 12. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal with all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ strictly larger than 1. Assume that $1 < \beta_1 \leq \beta_2 \leq \cdots \leq \beta_d$. For $1 \leq i \leq d$, let $\psi_i : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive function and $\mathbf{a} \in \mathbb{T}^d$. Assume that $\mathcal{U}(\Psi)$ is bounded. Then*

$$\dim_{\mathrm{H}} W(T, \Psi, \mathbf{a}) = \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\leq i\leq d} \big\{\theta_i(\mathbf{t})\big\},$$

where

$$\theta_i(\mathbf{t}) := \sum_{k\in\mathcal{K}_1(i)} 1 + \sum_{k\in\mathcal{K}_2(i)} \left(1 - \frac{t_k}{\log\beta_i + t_i}\right) + \sum_{k\in\mathcal{K}_3(i)} \frac{\log\beta_k}{\log\beta_i + t_i}$$

and, in turn

$$\mathcal{K}_1(i) := \{1 \leq k \leq d : \log\beta_k > \log\beta_i + t_i\},\ \mathcal{K}_2(i) := \{1 \leq k \leq d : \log\beta_k + t_k \leq \log\beta_i + t_i\},$$

and

$$\mathcal{K}_3(i) := \{1, \ldots, d\} \setminus (\mathcal{K}_1(i) \cup \mathcal{K}_2(i)).$$

Remark 11. It is easily seen the value of $\theta_i(\mathbf{t})$ remains unchanged if replace $>$ to $\geq$ in $\mathcal{K}_1(i)$, and/or replace $\leq$ by $<$ in $\mathcal{K}_2(i)$.

We now deduce Theorem 6 from Theorem 12.

Proof of Theorem 6 modulo Theorem 12. For the real positive function ψ in Theorem 6, we first suppose that its lower order at infinity is bounded; that is

$$\lambda = \lambda(\psi) := \liminf_{n\to\infty} \frac{-\log\psi(n)}{n} < +\infty.$$

With this in mind, put $\psi_1 = \psi_2 = \cdots = \psi_d = \psi$ in the statement of Theorem 12 and note that any $\mathbf{t}$ in $\mathcal{U}(\Psi)$ is of the form $\mathbf{t} = (t, t, \ldots, t)$ where $t \in \mathcal{U}(\psi)$ – the set of accumulation points of the sequence $\big\{-\frac{\log\psi(n)}{n}\big\}_{n\geq 1}$. We remark that $\lambda < +\infty$ means that $\mathcal{U}(\Psi)$ is bounded. Hence, for any $1 \leq i \leq d$ we have that

$$\theta_i(t) := \theta_i(\mathbf{t}) = \sum_{k:\beta_k>\beta_i e^t} 1 + \sum_{k:\beta_k\leq\beta_i}\left(1 - \frac{t}{\log\beta_i + t}\right) + \sum_{k:\beta_i e^t\geq\beta_k>\beta_i} \frac{\log\beta_k}{\log\beta_i + t}.$$

Now let

$$k_1 := \max\{1 \leq k \leq d : \beta_k \leq \beta_i\},$$

and

$$k_2 := \max\{1 \leq k \leq d : \beta_k \leq \beta_i e^t\}.$$

Then, the above expression for $\theta_i(t)$ becomes

$$\theta_i(t) = \sum_{k=k_2+1}^{d} 1 + \sum_{k=1}^{k_1} \frac{\log\beta_i}{\log\beta_i + t} + \sum_{k=k_1+1}^{k_2} \frac{\log\beta_k}{\log\beta_i + t} \tag{43}$$

and noting that $\beta_k = \beta_i$ for $i+1 \leq k \leq k_1$ whenever $k_1 > i$, it follows that

$$\begin{aligned}\theta_i(t) &= \frac{\sum_{k=k_2+1}^{d}(\log\beta_i + t) + i\log\beta_i + \sum_{k=i+1}^{k_2}\log\beta_k}{\log\beta_i + t} \\ &= \frac{i\log\beta_i - \sum_{k=k_2+1}^{d}(\log\beta_k - \log\beta_i - t) + \sum_{k=i+1}^{d}\log\beta_k}{\log\beta_i + t}.\end{aligned}$$

The upshot of this together with Theorem 12 is that

$$\dim_{\mathrm{H}} W(T, \psi, \mathbf{a}) = \sup_{t\in\mathcal{U}(\psi)} \min_{1\leq i\leq d} \theta_i(t).$$

Now observe that since $\theta_i(t)$ is a decreasing function in t, the above right hand side is equal to $\min_{1\leq i\leq d}\theta_i(\lambda)$ where $\lambda = \lambda(\psi)$ is the lower order at infinity of the function ψ. Thus, under the assumption that λ is bounded, we have that

$$\dim_{\mathrm{H}} W(T, \psi, \mathbf{a}) = \min_{1\leq i\leq d} \theta_i(\lambda)$$

as desired. To deal with the case that $\lambda = \lambda(\psi) = +\infty$, given any real number $M > 0$ consider the function $\psi_M : \mathbb{R}^+ \to \mathbb{R}^+ : x \to e^{-xM}$. Then, by definition $\mathcal{U}(\psi_M) = \{M\}$ and for M sufficiently large

$$W(T, \psi, \mathbf{a}) \subset W(T, \psi_M, \mathbf{a})$$

and so it follows that

$$0 \le \dim_{\mathrm{H}} W(T, \psi, \mathbf{a}) \le \dim_{\mathrm{H}} W(T, \psi_M, \mathbf{a}) \le \min_{1 \le i \le d} \theta_i(M). \tag{44}$$

Now with reference to (43), we have that $t = M$ and so for M sufficiently large: $k_2 = d$ for $1 \le i \le d$. Hence, for any $1 \le i \le d$

$$\lim_{M \to \infty} \theta_i(M) = \lim_{M \to \infty} \left(\sum_{k=1}^{i} \frac{\log \beta_i}{\log \beta_i + M} + \sum_{k=i+1}^{d} \frac{\log \beta_k}{\log \beta_i + M} \right) = 0.$$

This together with (44) implies that

$$\dim_{\mathrm{H}} W(T, \psi, \mathbf{a}) = 0. \quad \square$$

4.2.1. Proof of Theorem 12

We start with a brief discussion that sums up various fundamental notions and statements that we will require during the course of establishing Theorem 12. The statements are concerned with the distribution of the preimages of a fixed ball under a given β-transformation T_β. As usual, let $\beta \in \mathbb{R}$ such that $|\beta| > 1$ and let

$$\mathcal{Q} = \left\{ \Big[0, \frac{1}{|\beta|}\Big), \dots, \Big[\frac{k}{|\beta|}, \frac{k+1}{|\beta|}\Big), \dots, \Big[\frac{\lfloor|\beta|\rfloor}{|\beta|}, 1\Big) \right\}$$

be the natural partition of $[0,1)$. The n-th refinement of $\mathcal{Q}$ is defined as

$$\mathcal{Q}^n := \left\{ Q_{i_0} \cap T_\beta^{-1}(Q_{i_1}) \cap \dots \cap T_\beta^{-(n-1)}(Q_{i_{n-1}}) : \quad Q_{i_j} \in \mathcal{Q} \text{ for } 0 \le j \le n-1 \right\}.$$

The elements in $\mathcal{Q}^n$ are called *cylinders of order* n. Evidently, the cylinders are disjoint and the restriction of T_β^n on each cylinder is continuous linear of slop β^n. Now given a point $a \in \mathbb{T}$, consider the preimage of the ball $B(a, r)$ under T_β^n. It can be verified that this preimage consists of disjoint intervals whose lengths are bounded above by $2r|\beta|^{-n}$. Indeed, we can write

$$T_\beta^{-n}\big(B(a, r)\big) = \bigcup_{j=1}^{N_n} I_{n,j}, \tag{45}$$

where each $I_{n,j}$ is an interval lying in some cylinder of order n and N_n is the number of such intervals.

For $\beta > 1$, via the work of Rényi [39], it follows that the (total) number of cylinders of order n of T_β is bounded from above by $\beta^{n+1}/(\beta-1)$. Hence,

$$N_n \le \beta^{n+1}/(\beta-1). \tag{46}$$

Remark 12. It is worth mentioning that on exploiting the well-known fact that the topological entropy of T_β is $\log|\beta|$ for any β with $|\beta|>1$, we obtain the weaker bound

$$N_n \le |\beta|^{n(1+\epsilon)} \tag{47}$$

for any $\epsilon > 0$ and n sufficiently large. This suffices for not only proving Lemma 8 below but more importantly for establishing the upper bound for the dimension within the context of Theorems 7 and 9 in which β is allowed to be negative.

Now, suppose $\beta > 1$. Recall that a cylinder I of order n is said to be *full* for T_β if $T_\beta^n(I) = \mathbb{T}$. With this in mind, Bugeaud & Wang [8, Theorem 1.2] proved that every $(n+1)$ consecutive cylinders of order n contains at least one full cylinder. This gives rise to the following useful fact that we will make use of on multiple occasions.

Fact BW: *The distance between any two consecutive full cylinders is less than $(n+1)\beta^{-n}$. Furthermore, since any full cylinder intersects $T_\beta^{-n}\big(B(a,r)\big)$ it follows that the distance between any two consecutive intervals $I_{n,j}$ and $I_{n,j+1}$ is less than $(n+3)\beta^{-n}$; i.e.,*

$$\operatorname{dist}(I_{n,j}, I_{n,j+1}) \le (n+3)\beta^{-n}.$$

We now move onto the task of proving Theorem 12. This will be done by establishing the upper and lower bounds for $\dim W(T,\Psi,\mathbf{a})$ separately.

Proposition 4. *Under the setting of Theorem 12, we have that*

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \le \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\le i\le d}\{\theta_i(\mathbf{t})\}.$$

We first establish the proposition in the special case that $\mathcal{U}(\Psi)$ consists of a single point.

Lemma 8. *Under the setting of Theorem 12, assume in addition that there exists $\mathbf{t} = (t_1,\dots,t_d) \in (\mathbb{R}^+)^d$ such that*

$$\lim_{n\to\infty} \frac{-\log\psi_i(n)}{n} = t_i \quad \text{for all} \quad 1\le i\le d.$$

Then

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \le \min_{1\le i\le d}\{\theta_i(\mathbf{t})\}.$$

Proof. Observe that we can re-write (42) as

$$W(T,\Psi,\mathbf{a}) = \limsup_{n\to\infty}\ T_{\beta_1}^{-n}\big(B(a_1,\psi_1(n))\big)\times\cdots\times T_{\beta_d}^{-n}\big(B(a_d,\psi_d(n))\big), \tag{48}$$

where T_{β_i} is the standard β-transformation with $\beta=\beta_i$. As usual, we do not distinguish between β-transformations acting on the unit interval $[0,1)$ or the torus $\mathbb{T}$. The proof of Lemma 8 relies on finding an "efficient" covering by balls of the lim sup set (48). So with this in mind, for any $1\le i\le d$, by (45) we have that

$$T_{\beta_i}^{-n}\big(B(a_i,\psi_i(n))\big) = \bigcup_{j=1}^{N_{i,n}} I_{n,j}^{(i)}\,, \tag{49}$$

where each $I_{n,j}^{(i)}$ is an interval lying in some cylinder of order n and $N_{i,n}$ is the number of such intervals. Now for $n\in\mathbb{N}$, let

$$J_n := \big\{\mathbf{j}=(j_1,\ldots,j_d) : 1\le j_i\le N_{i,n}\ (1\le i\le d)\big\}$$

and for $\mathbf{j}\in J_n$, let

$$R_{n,\mathbf{j}} := I_{n,j_1}^{(1)}\times\cdots\times I_{n,j_d}^{(d)}\,.$$

In turn, let

$$R_n = \bigcup_{\mathbf{j}\in J_n} R_{n,\mathbf{j}} \qquad\text{and}\qquad \mathcal{R}_n := \big\{R_{n,\mathbf{j}}\ :\ \mathbf{j}\in J_n\big\}\,.$$

Then in view of (49), we can re-write (48) as

$$W(T,\Psi,\mathbf{a}) = \bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty} R_n$$

and it follows that for any $N\ge 1$,

$$W(T,\Psi,\mathbf{a}) \subset \bigcup_{n=N}^{\infty} R_n\,.$$

In other words, the collection $\{\mathcal{R}_n : n=N,N+1,\ldots\}$ of rectangles $R_{n,\mathbf{j}}$ form a cover for the set $W(T,\Psi,\mathbf{a})$. Now, observe that along the direction of the i-th axis $(1\le i\le d)$, by construction for each $1\le j<N_{i,n}$ the sides $I_{n,j}^{(i)}$ and $I_{n,j+1}^{(i)}$ are disjoint and thus the rectangles in $\mathcal{R}_n$ are disjoint. On the other hand, by Fact BW

$$\mathrm{dist}(I_{n,j}^{(i)}, I_{n,j+1}^{(i)}) \le (n+3)\beta_i^{-n}$$

and so along the direction of the i-th axis, the distance between consecutive rectangles in $\mathcal{R}_n$ is at most $(n+3)\beta_i^{-n}$.

We now estimate the number of balls $B_{i,n}$ of diameter $2\psi_i(n)\beta_i^{-n}$ (the sidelength of the rectangles in $\mathcal{R}_n$ along the direction of the i-th axis) needed to cover the set R_n. We start by covering a fixed generic rectangle $R = R_{n,\mathbf{j}} \in \mathcal{R}_n$. It is easily verified that we can find a collection $\mathcal{B}_{i,n}(R)$ of balls $B_{i,n}$ that covers R with

$$\#\mathcal{B}_{i,n}(R) \;\le\; 2^d \prod_{\substack{1\le k\le d\,:\\ \psi_k(n)\beta_k^{-n}\ge\psi_i(n)\beta_i^{-n}}} \frac{\psi_k(n)\beta_k^{-n}}{\psi_i(n)\beta_i^{-n}}\,. \tag{50}$$

Indeed, we can simply take the natural cover in which we split R into closed balls $B_{i,n}$ which are disjoint apart from at the boundary. Now observe that the collection $\mathcal{B}_{i,n}(R)$ will also cover other rectangles in $\mathcal{R}_n$ along the direction of the k-th axis ($1 \le k \le d$) if the separation in that direction is small compared to the diameter of the balls $B_{i,n}$; that is, in view of Fact BW if

$$(n+3)\beta_k^{-n} < 2\psi_i(n)\beta_i^{-n}\,.$$

In particular, this leads to the following lower bound for the number $M_{i,n}(R)$ of rectangles covered by $\mathcal{B}_{i,n}(R)$:

$$\begin{aligned} M_{i,n}(R) &:= \#\Big\{R_{n,\mathbf{j}} \in \mathcal{R}_n \;:\; R_{n,\mathbf{j}} \subseteq \bigcup_{B_{i,n}\in\mathcal{B}_{n,i}(R)} B_{i,n}\Big\} \\ &\ge \prod_{\substack{1\le k\le d\,:\\ (n+3)\beta_k^{-n}<\psi_i(n)\beta_i^{-n}}} \frac{2\psi_i(n)\beta_i^{-n}}{(n+3)\beta_k^{-n}}\,. \end{aligned} \tag{51}$$

The upshot is that there is a collection $\mathcal{B}_{n,i}$ of balls of $B_{i,n}$ that cover the set R_n with

$$\#\mathcal{B}_{n,i} \;\le\; \frac{\#\mathcal{R}_n}{M_{i,n}(R)}\;\#\mathcal{B}_{n,i}(R) \;=\; \prod_{j=1}^{d} N_{j,n}\;\cdot\;\frac{\#\mathcal{B}_{n,i}(R)}{M_{i,n}(R)}\,.$$

This together with (46), (50) and (51) implies that

$$\begin{aligned} \#\mathcal{B}_{n,i} \;&\le \prod_{j=1}^{d}\frac{\beta_j^{n+1}}{\beta_j-1}\cdot \prod_{\substack{1\le k\le d\,:\\ (n+3)\beta_k^{-n}<\psi_i(n)\beta_i^{-n}}} \frac{(n+3)\beta_k^{-n}}{2\psi_i(n)\beta_i^{-n}} \\ &\quad\cdot\; 2^d \prod_{\substack{1\le k\le d\,:\\ \psi_k(n)\beta_k^{-n}\ge\psi_i(n)\beta_i^{-n}}} \frac{\psi_k(n)\beta_k^{-n}}{\psi_i(n)\beta_i^{-n}} \end{aligned}$$

$$= 2^d \cdot \prod_{j=1}^{d} \frac{\beta_j^{n+1}}{\beta_j - 1} \prod_{k \in \mathcal{K}_{n,1}(i)} \frac{(n+3)\beta_k^{-n}}{2\psi_i(n)\beta_i^{-n}} \prod_{k \in \mathcal{K}_{n,2}(i)} \frac{\psi_k(n)\beta_k^{-n}}{\psi_i(n)\beta_i^{-n}},$$

where

$$\begin{aligned}\mathcal{K}_{n,1}(i) &:= \left\{1 \le k \le d : (n+3)\beta_k^{-n} < \psi_i(n)\beta_i^{-n}\right\} \\ &= \left\{1 \le k \le d : -\frac{\log(n+3)}{n} + \log\beta_k > -\frac{\log\psi_i(n)}{n} + \log\beta_i\right\},\end{aligned}$$

and

$$\begin{aligned}\mathcal{K}_{n,2}(i) &:= \left\{1 \le k \le d : \psi_k(n)\beta_k^{-n} \ge \psi_i(n)\beta_i^{-n}\right\} \\ &= \left\{1 \le k \le d : -\frac{\log\psi_k(n)}{n} + \log\beta_k \le -\frac{\log\psi_i(n)}{n} + \log\beta_i\right\}.\end{aligned}$$

Thus, given $\rho > 0$ and on choosing N sufficiently large so that $2\psi_i(n)\beta_i^{-n} < \rho$ for any $n \ge N$, it follows from the definition of s-dimensional Hausdorff measure that for any $s > 0$

$$\begin{aligned}&\mathcal{H}_\rho^s(W(T, \Psi, \mathbf{a})) \\ &\le \sum_{n=N}^{\infty} \#\mathcal{B}_{n,i} \, \left(2\psi_i(n)\beta_i^{-n}\right)^s \\ &\le \sum_{n=N}^{\infty} 2^d \cdot \prod_{j=1}^{d} \frac{\beta_j^{n+1}}{\beta_j - 1} \prod_{k \in \mathcal{K}_{n,1}(i)} \frac{(n+3)\beta_k^{-n}}{2\psi_i(n)\beta_i^{-n}} \prod_{k \in \mathcal{K}_{n,2}(i)} \frac{\psi_k(n)\beta_k^{-n}}{\psi_i(n)\beta_i^{-n}} \cdot \left(2\psi_i(n)\beta_i^{-n}\right)^s \\ &= C \sum_{n=N}^{\infty} \exp\left\{-n \cdot \ell_n\right\},\end{aligned} \tag{52}$$

where $C := 2^{s+d} \prod_{j=1}^{d} \frac{\beta_j}{\beta_j - 1}$ is a constant and

$$\begin{aligned}\ell_n = \ell_n(i) \;:=\; -\sum_{j=1}^{d} \log\beta_j \quad &- \sum_{k \in \mathcal{K}_{n,1}(i)} \left(\frac{\log(n+3)}{n} - \log\beta_k - \frac{\log\psi_i(n)}{n} + \log\beta_i\right) \\ &- \sum_{k \in \mathcal{K}_{n,2}(i)} \left(\frac{\log\psi_k(n)}{n} - \log\beta_k - \frac{\log\psi_i(n)}{n} + \log\beta_i\right) \\ &+ \quad s\left(-\frac{\log\psi_i(n)}{n} + \log\beta_i\right).\end{aligned}$$

Now note that $\sum_{n=1}^{\infty} \exp\left\{-n \cdot \ell_n\right\}$ converges as long as

$$\limsup_{n \to \infty} \ell_n > 0,$$

and that this is equivalent to the condition that s is strictly larger than the upper limit of

$$\begin{aligned}
h_n = h_n(i) \; := \; & \frac{\sum_{j=1}^{d} \log\beta_j + \sum_{k\in\mathcal{K}_{n,1}(i)} \left(\frac{\log(n+3)}{n} - \log\beta_k - \frac{\log\psi_i(n)}{n} + \log\beta_i\right)}{-\frac{\log\psi_i(n)}{n} + \log\beta_i} \\
& + \quad \frac{\sum_{k\in\mathcal{K}_{n,2}(i)} \left(\frac{\log\psi_k(n)}{n} - \log\beta_k - \frac{\log\psi_i(n)}{n} + \log\beta_i\right)}{-\frac{\log\psi_i(n)}{n} + \log\beta_i} \\
= & \sum_{k\in\mathcal{K}_{n,1}(i)} 1 + \sum_{k\in\mathcal{K}_{n,2}(i)} \left(1 - \frac{-\frac{\log\psi_k(n)}{n}}{-\frac{\log\psi_i(n)}{n} + \log\beta_i}\right) + \sum_{k\in\mathcal{K}_{n,3}(i)} \frac{\log\beta_k}{-\frac{\log\psi_i(n)}{n} + \log\beta_i},
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{K}_{n,3}(i) &:= \{1,\ldots,d\} \setminus (\mathcal{K}_{n,1}(i) \cup \mathcal{K}_{n,2}(i)) \\
&= \left\{ 1 \le k \le d : -\frac{\log(n+3)}{n} + \log\beta_k \right. \\
&\quad \left. \le -\frac{\log\psi_i(n)}{n} + \log\beta_i < -\frac{\log\psi_k(n)}{n} + \log\beta_k \right\}.
\end{aligned}$$

So, by the additional assumption imposed in the lemma, it follows that

$$\begin{aligned}
\limsup_{n\to\infty} h_n = \lim_{n\to\infty} h_n = & \sum_{k\in\mathcal{K}_1(i)} 1 + \sum_{k\in\mathcal{K}_2(i)} \left(1 - \frac{t_k}{\log\beta_i + t_i}\right) \\
& + \sum_{k\in\mathcal{K}_3(i)} \frac{\log\beta_k}{\log\beta_i + t_i} = \theta_i(\mathbf{t}).
\end{aligned}$$

The upshot of the above is that for any $1 \le i \le d$ and $s > \theta_i(\mathbf{t})$, we have that

$$\sum_{n=1}^{\infty} \exp\{-n\cdot\ell_n\} < \infty$$

and hence together with (52) we obtain that

$$0 \le \mathcal{H}^s(W(T,\Psi,\mathbf{a})) = \lim_{\rho\to 0} \mathcal{H}^s_\rho(W(T,\Psi,\mathbf{a})) \le \lim_{N\to\infty} C \sum_{n=N}^{\infty} \exp\{-n\cdot\ell_n\} = 0\,.$$

In turn, it follows from the definition of Hausdorff dimension that $\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \le \theta_i(\mathbf{t})$. This upper bound estimate is true for any $1 \le i \le d$, and so it implies that

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \le \min_{1\le i\le d} \theta_i(\mathbf{t})$$

as desired. □

Armed with Lemma 8, it is relatively straightforward to prove the general upper bound statement for the Hausdorff dimension of $W(T,\Psi,\mathbf{a})$.

Proof of Proposition 4. To prove the proposition, we first cover the accumulation set $\mathcal{U}(\Psi)$. For any $\varepsilon>0$, since $\mathcal{U}(\Psi)$ is bounded, we can find a family $\mathcal{B}_\epsilon$ of finitely many balls of the form

$$B=\prod_{i=1}^{d}[b_B^{(i)},b_B^{(i)}+\varepsilon]\qquad (b_B^{(i)}\geq 0)$$

that cover $\mathcal{U}(\Psi)$. For $B\in\mathcal{B}_\epsilon$, let

$$\mathcal{N}(B)=\left\{n\in\mathbb{N}:\left(\frac{-\log\psi_1(n)}{n},\cdots,\frac{-\log\psi_d(n)}{n}\right)\in\prod_{i=1}^{d}[b_B^{(i)},b_B^{(i)}+\varepsilon]\right\}.$$

Without loss of generality, we assume that $\#\mathcal{N}(B)=\infty$ since, otherwise, there is no accumulation point in the ball B. We claim that $W(T,\Psi,\mathbf{a})$ is a subset of

$$\bigcup_{B\in\mathcal{B}}\left\{\mathbf{x}\in\mathbb{T}^d:\|T_{\beta_i}^n x_i-a_i\|\leq e^{-nb_B^{(i)}}\ \ (1\leq i\leq d)\ \text{ for infinitely many } n\in\mathcal{N}(B)\right\}.$$

Indeed, for any $\mathbf{x}\in W(T,\Psi,\mathbf{a})$, there exists a sequence $\{n_j\}_{j\in\mathbb{N}}$ depending on $\mathbf{x}$ such that for any $1\leq i\leq d$

$$\|T_{\beta_i}^{n_j}x_i-a_i\|\leq\psi_i(n_j)\quad\forall\ \ j\geq 1.$$

Since there are only finitely many balls $B\in\mathcal{B}_\epsilon$ which cover $\mathcal{U}(\Psi)$, there exists some $B\in\mathcal{B}_\epsilon$ that contains infinitely many points of

$$\left\{\Big(\frac{-\log\psi_1(n_j)}{n_j},\cdots,\frac{-\log\psi_d(n_j)}{n_j}\Big)\right\}_{j\in\mathbb{N}}.$$

Thus, for these infinitely many j's, we have that for any $1\leq i\leq d$

$$\|T_{\beta_i}^{n_j}x_i-a_i\|\ \leq\ \psi_i(n_j)\ \leq\ e^{-n_j b_B^{(i)}}.$$

This establishes the claim and by the countable stability property of Hausdorff dimension, it follows that $\dim_{\mathrm{H}}W(T,\Psi,\mathbf{a})$ is less than or equal to

$$\max_{B\in\mathcal{B}_\epsilon}\ \dim_{\mathrm{H}}\left\{\mathbf{x}\in\mathbb{T}^d:\|T_{\beta_i}^n x_i-a_i\|\leq e^{-nb_B^{(i)}}\ (1\leq i\leq d)\ \text{ for infinitely many } n\in\mathbb{N}\right\}.$$

Now observe that

$$\lim_{n\to\infty} \frac{-\log e^{-nb_B^{(i)}}}{n} \quad \text{for all} \quad 1 \le i \le d\,,$$

and so on applying Lemma 8 we deduce that

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \le \max_{B\in\mathcal{B}_\epsilon} \min_{1\le i\le d} \theta_i\big((b_B^{(1)},\ldots,b_B^{(d)})\big).$$

Then on letting $\varepsilon \to 0$, by the continuity of $\theta_i(\mathbf{t})$ with respect to $\mathbf{t}$, we conclude that

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \le \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\le i\le d} \{\theta_i(\mathbf{t})\}.$$

This completes the proof of Proposition 4. $\square$

We now turn out attention to establishing the lower bound for the Hausdorff dimension of $W(T,\Psi,\mathbf{a})$.

Proposition 5. *Under the setting of Theorem 12, we have that*

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \ge \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\le i\le d} \{\theta_i(\mathbf{t})\}.$$

Proof. The proof of Proposition 5 relies on constructing a suitable lim sup type subset of $W(T,\Psi,\mathbf{a})$ which enables us to exploit the 'rectangles to rectangles' Mass Transference Principle (Theorem 11). With this in mind, by (48) and (49), we know that $W(T,\Psi,\mathbf{a})$ is a limsup set of rectangles with sides given by the intervals $I_{n,j}^{(i)}$ $(1\le i\le d)$. Recall, that the sides correspond to the intersection of $T_{\beta_i}^{-n}\big(B(a_i,\psi_i(n))\big)$ and a cylinder of order n and as in the proof of Proposition 4, we do not distinguish between β-transformations acting on the unit interval or the torus. Thus, if for each $1\le i\le d$, we select only those intervals which are intersections of $T_{\beta_i}^{-n}\big(B(a_i,\psi_i(n))\big)$ and full cylinders of order n, we will obtain a lim sup type subset of $W(T,\Psi,\mathbf{a})$; that is

$$W(T,\Psi,\mathbf{a}) \supset \bigcap_{N=1}^{\infty}\bigcup_{n=N}^{\infty}\bigcup_{j_1=1}^{M_{1,n}}\cdots\bigcup_{j_d=1}^{M_{d,n}} B(x_{n,j_1}^{(1)},\beta_1^{-n}\psi_1(n))\times\cdots\times B(x_{n,j_d}^{(d)},\beta_d^{-n}\psi_d(n))\,, \quad (53)$$

where $\{x_{n,j_i}^{(i)}, 1\le j_i\le M_{i,n}\}$ are the preimages of a_i under $T_{\beta_i}^n$ that fall within full cylinders of order n for T_{β_i} and $M_{i,n}$ is the number of such full cylinders. Now with (53) and Fact BW in mind, it follows that for each $1\le i\le d$ the enlarged collection of balls or rather intervals $\big\{B(x_{n,j_i}^{(i)},(n+3)\beta_i^{-n}) : 1\le j_i\le M_{i,n}\big\}$ covers $\mathbb{T}$, that is

$$\mathbb{T} \;=\; \bigcup_{j_i=1}^{M_{i,n}} B\big(x_{n,j_i}^{(i)},(n+3)\beta_i^{-n}\big)\,. \quad (54)$$

Now fix a point $\mathbf{t} = (t_1, \dots, t_d) \in \mathcal{U}(\Psi)$. Then by definition and the fact that $\mathcal{U}(\Psi)$ is bounded, there exists a subsequence $\{n_l\}_{l\in\mathbb{N}}$ such that

$$\lim_{l\to\infty} \frac{-\log\psi_i(n_l)}{n_l} = t_i \quad \text{for all} \quad 1 \le i \le d.$$

It is easily verified that for any $0 < \varepsilon < 1$, there exists $N = N(\varepsilon) > 0$ such that

$$\frac{(1-\varepsilon)\log\beta_i}{(1-\varepsilon)\log\beta_i + t_i} \le \frac{-\frac{\log(n_l+3)}{n_l} + \log\beta_i}{\log\beta_i + \frac{-\log\psi_i(n_l)}{n_l}} \tag{55}$$

for all $l \ge N$ and $1 \le i \le d$. Let

$$s_i := \frac{(1-\varepsilon)\log\beta_i}{(1-\varepsilon)\log\beta_i + t_i} \qquad (1 \le i \le d).$$

Then, (55) is equivalent to

$$\big(\beta_i^{-n_l}\psi_i(n_l)\big)^{s_i} \ge (n_l+3)\beta_i^{-n_l},$$

which together with (54) implies that for any $l \ge N$

$$\mathbb{T} = \bigcup_{j_i=1}^{M_{i,n_l}} B\big(x^{(i)}_{n_l,j_i}, \big(\beta_i^{-n_l}\psi_i(n_l)\big)^{s_i}\big).$$

In turn, it follows that for any $l \ge N$

$$\mathbb{T}^d = \bigcup_{j_1=1}^{M_{1,n_l}} \cdots \bigcup_{j_d=1}^{M_{d,n_l}} B\big(x^{(1)}_{n_l,j_1}, (\beta_1^{-n_l}\psi_1(n_l))^{s_1}\big) \times \cdots \times B\big(x^{(d)}_{n_l,j_d}, (\beta_d^{-n_l}\psi_d(n_l))^{s_d}\big)$$

and so

$$\mathbb{T}^d = \limsup_{n\to\infty} \bigcup_{j_1=1}^{M_{1,n}} \cdots \bigcup_{j_d=1}^{M_{d,n}} B\big(x^{(1)}_{n,j_1}, (\beta_1^{-n}\psi_1(n))^{s_1}\big) \times \cdots \times B\big(x^{(d)}_{n,j_d}, (\beta_d^{-n}\psi_d(n))^{s_d}\big). \tag{56}$$

The upshot is that given the lim sup set of rectangles appearing on the right hand side of (53), the corresponding lim sup set of '$(s_1, \dots, s_d)$-scaled up' rectangles satisfies (41) with $p = d$, $X_i = \mathbb{T}$, $\delta_i = 1$ and $\mu_i = m_1$ (one-dimensional Lebesgue measure) for each $1 \le i \le d$. Thus on applying Theorem 11 with $u_i = (1-\varepsilon)\log\beta_i$ and $v_i = (1-\varepsilon)\log\beta_i + t_i$ $(1 \le i \le d)$, we obtain the lower bound

$$\dim_{\mathrm{H}} W(T, \Psi, \mathbf{a}) \ge \min_{1\le i\le d} s(i, \varepsilon)$$

where

$$s(i,\varepsilon) := \sum_{k\in\mathcal{K}_1(i,\varepsilon)} 1 + \sum_{k\in\mathcal{K}_2(i,\varepsilon)} \left(1 - \frac{t_k}{(1-\varepsilon)\log\beta_i + t_i}\right) + \sum_{k\in\mathcal{K}_3(i,\varepsilon)} \frac{(1-\varepsilon)\log\beta_k}{(1-\varepsilon)\log\beta_i + t_i}$$

and where $\mathcal{K}_1(i,\varepsilon)$, $\mathcal{K}_2(i,\varepsilon)$, $\mathcal{K}_3(i,\varepsilon)$ is the partition of $\{1,\dots,d\}$ given by

$$\begin{aligned}
\mathcal{K}_1(i,\varepsilon) &:= \big\{k : (1-\varepsilon)\log\beta_k \geq (1-\varepsilon)\log\beta_i + t_i\big\}\\
\mathcal{K}_2(i,\varepsilon) &:= \big\{k : (1-\varepsilon)\log\beta_k + t_k \leq (1-\varepsilon)\log\beta_i + t_i\big\},\\
\mathcal{K}_3(i,\varepsilon) &:= \big\{1,\dots,d\big\} \setminus \big(\mathcal{K}_1(i,\varepsilon)\cup\mathcal{K}_2(i,\varepsilon)\big).
\end{aligned}$$

Fix $1 \leq i \leq d$. On letting $\varepsilon \to 0$, we find that $\mathcal{K}_1(i,\varepsilon) \to \{k : \log\beta_k > \log\beta_i + t_i\} = \mathcal{K}_1(i)$, $\mathcal{K}_2(i,\varepsilon) \to \mathcal{K}_2(i)$ and $\mathcal{K}_3(i,\varepsilon) \to \mathcal{K}_3(i)$. Thus

$$\lim_{\varepsilon\to 0} s(i,\varepsilon) = \theta_i(\mathbf{t}) \quad \text{and} \quad \dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \geq \min_{1\leq i\leq d} \theta_i(\mathbf{t}).$$

Moreover, since this is valid for any $\mathbf{t} \in \mathcal{U}(\Psi)$ it follows that

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \geq \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\leq i\leq d} \{\theta_i(\mathbf{t})\}$$

and we are done. $\square$

4.2.2. Proof of Theorem 8

We show that when T is an integer matrix transformation, the diagonal assumption in Theorem 6 can be relaxed to T is diagonalizable over $\mathbb{Z}$. This thereby proves Theorem 8. So, suppose T is diagonalizable over $\mathbb{Z}$. Then by definition, there exist a diagonal integer matrix D and an invertible mapping ϕ satisfying (37). It is easily versified that $T^n(\mathbf{x}) \in B(\mathbf{a},\psi(n))$ if and only if $D^n(\phi(\mathbf{x})) \in \phi\big(B(\mathbf{a},\psi(n))\big)$. Since ϕ is a bi-Lipschitz map, we can find two positive constants $0 < c_1 \leq c_2 < \infty$ such that

$$B\big(\phi(\mathbf{a}), c_1\psi(n)\big) \subset \phi\big(B(\mathbf{a},\psi(n)\big) \subset B\big(\phi(\mathbf{a}), c_2\psi(n)\big).$$

In turn, Lemma 7 implies that the Hausdorff dimension of

$$W(T,\psi,\mathbf{a}) := \{\mathbf{x}\in\mathbb{T}^d : T^n(\mathbf{x}) \in B(\mathbf{a},\psi(n)) \ \text{ for infinitely many } n\in\mathbb{N}\}$$

is the same as that of

$$\big\{x\in\mathbb{T}^d : D^n(\mathbf{x}) \in B(\phi(\mathbf{a}),\psi(n)) \ \text{ for infinitely many } n\in\mathbb{N}\big\}.$$

Thus, without loss of generality, we only need to prove the desired dimension result in the case that T is diagonal.

4.3. Proof of Theorem 7

The proof of Theorem 7 will make use of a general statement (namely, Proposition 7 below) concerning Markov subsystems which may be of independent interest. In short, these systems provide a "nice" approximation to one-dimensional piecewise linear dynamical systems. To start with, let us recall the notion of a Markov system for a one-dimensional expanding dynamical system (X, T). With this in mind, let X be a compact set in $\mathbb{R}$ and $T : X \to X$ be an expanding map. Furthermore, let Λ be a subset of X. A partition $\mathcal{P}_\Lambda$ of Λ into finite or countable collection of sets $P(k)$ is called a *Markov partition* if $\Lambda := \bigcap_{n=0}^{\infty} T^{-n}\big(\cup P(k)\big)$ and

(i) the interior of $P(j)$ and $P(k)$ are disjoint if $j \neq k$,
(ii) T restricted on each $P(j)$ is one to one,
(iii) if $T(P(j))$ intersects the interior of $P(k)$ for some j and k then $P(k) \subseteq \overline{T(P(j))}$.

In turn, the system $(\Lambda, T|_\Lambda, \mathcal{P}_\Lambda)$ is called a *Markov subsystem of* (X, T). In the case $\Lambda = X$, we simply write $(X, T, \mathcal{P})$ and referred to it as a *Markov system*.

An important property regarding Markov subsystems that we shall utilise is given by the following statement. It is a direct consequence of [34, Theorems 4.2.9 & 4.2.11].

Proposition 6. *Let X be a compact set in $\mathbb{R}$ and $T : X \to X$ be an expanding map. Let $(\Lambda, T|_\Lambda, \mathcal{P}_\Lambda)$ be a Markov subsystem of (X, T) with finite partition $\mathcal{P}_\Lambda = \{P(i)\}_{1 \le i \le N}$ whose incidence matrix is primitive. Suppose that for any $1 \le k \le N$, $T|_{P(k)}$ is $C^{1+\alpha}$ for some $\alpha > 0$. Then the measure $\mathcal{H}^\delta|_\Lambda$ is δ-Ahlfors regular where $\delta := \dim_{\mathrm{H}} \Lambda$.*

The following statement provides a lower bound for $\dim_{\mathrm{H}} \Lambda$ in the case T is piecewise linear. Throughout, we suppose that the absolute value of the slope of such a map T is constant and will be denoted by $\beta(T)$.

Proposition 7. *Let T be a piecewise linear map on $[0, 1]$ and assume that $\beta(T) > 8$. Then there exists a Markov subsystem $(\Lambda, T|_\Lambda, \mathcal{P}_\Lambda)$ of $([0, 1], T)$ with a finite partition $\mathcal{P}_\Lambda = \{P(i)\}_{1 \le i \le m}$ where each $P(i)$ is an interval and $T|_{P(i)}$ is linear, such that*

$$\dim_{\mathrm{H}} \Lambda \ge 1 - \frac{\log 8}{\log \beta(T)}.$$

Proof. Let $\tilde{\mathcal{P}} = \{\tilde{P}(i)\}_{i=1}^m$ be a partition of $[0, 1]$ such that for each $1 \le i \le m$ the set $\tilde{P}(i)$ is an interval and $T|_{\tilde{P}(i)}$ is linear. Without loss of generality, we can assume that

$$\max\{|\tilde{P}(i)| : \tilde{P}(i) \in \tilde{\mathcal{P}}\} \le 2\kappa \quad \text{with} \quad \kappa := \min\{|\tilde{P}(i)| : \tilde{P}(i) \in \tilde{\mathcal{P}}\}.$$

Indeed, if $|\tilde{P}(i)| > 2\kappa$ for some $1 \le i \le m$, then there exists $\ell \in \mathbb{N}$ such that

$$2^{\ell}\kappa < |\tilde{P}(i)| \leq 2^{\ell+1}\kappa.$$

Hence, we can subdivide $\tilde{P}(i)$ into 2^{ℓ} equal pieces and take these subintervals as part of partition rather than $\tilde{P}(i)$. The map T restricted to each piece of the new partition is still linear and by construction the length of each piece if bounded above by 2κ.

For any interval $\tilde{P} \in \tilde{\mathcal{P}}$, let

$$P := \tilde{P} \cap T^{-1}\Big(\overline{\bigcup_{1\leq i\leq m} \{\tilde{P}(i) \in \tilde{\mathcal{P}} : \tilde{P}(i) \subset T(\tilde{P})\}} \Big).$$

Now since $T|_{\tilde{P}(i)}$ is linear, the intervals $\tilde{P}(i)$ contained in $T(\tilde{P})$ are adjacent intervals in the partition $\tilde{\mathcal{P}}$. Hence, P is a subinterval of $\tilde{P}$. Furthermore, since $T|_{\tilde{P}}$ is linear with slope $\pm\beta(T)$ we have that $|T(\tilde{P})| = \beta(T)\cdot|\tilde{P}| \geq \beta(T)\kappa$. So the number of $\tilde{P}(i) \in \tilde{\mathcal{P}}$ that intersect $T(\tilde{P})$ is at least the integer part of $\beta(T)\kappa/2\kappa = [\beta(T)/2]$. Here we use the fact that $|\tilde{P}(i)| \leq 2\kappa$ for all intervals in the partition. Thus, on using the fact that $\beta(T) > 8$, we have that

$$\#\big\{\tilde{P}(i) \in \tilde{\mathcal{P}} : \tilde{P}(i) \subset T(\tilde{P})\big\} \geq [\beta(T)/2] - 2 \geq 1\,. \tag{57}$$

The upshot of this is that

$$P \neq \emptyset\,.$$

We now prove that $\mathcal{P} := \{P(i) : \tilde{P}(i) \in \tilde{\mathcal{P}},\ 1 \leq i \leq m\}$ is a Markov partition of $\bigcup_{i=1}^{m} P(i)$. The first two conditions are automatically satisfied. Regarding the third condition, for any $1 \leq j,k \leq m$ with $T(P(j))\cap P(k) \neq \emptyset$, we first note that $T(P(j))\cap\tilde{P}(k) \neq \emptyset$ which in turn implies that $P(j)\cap T^{-1}(\tilde{P}(k)) \neq \emptyset$. Then, by the definition of $P(j)$, $\tilde{P}(k)$ is an interval such that $\tilde{P}(k) \subset T(\tilde{P}(j))$ and $\tilde{P}(j)\cap T^{-1}(\overline{\tilde{P}(k)}) \subset P(j)$. So $\overline{\tilde{P}(k)} \subset T(P(j))$. Therefore, by noting that $P(k) \subset \tilde{P}(k)$, we have $P(k) \subset T(P(j)) \subset \overline{T(P(j))}$ and this verifies the third condition.

Next, let

$$f := T|_{\cup_{i=1}^{m} P(i)} \quad \text{and} \quad \Lambda := \bigcap_{n=0}^{\infty} f^{-n}\big(\cup_{i=1}^{m} P(i)\big)\,.$$

Then, by construction, the system $(\Lambda, T|_\Lambda, \mathcal{P}_\Lambda)$ with $\mathcal{P}_\Lambda := \{P(i) : \tilde{P}(i) \in \mathcal{P},\ 1 \leq i \leq m\}$ is a Markov subsystem of $([0,1], T)$. It remains to prove that the Hausdorff dimension of the set Λ satisfies the lower bound in the statement of the proposition. For this, we work in the symbolic space of the dynamical system under consideration to estimate the topological entropy of $T|_\Lambda$ and then use the fact that the entropy is intimately related to the dimension of Λ.

The dynamics of $T|_\Lambda$ can be coded by the $m\times m$ transition matrix $A = (A_{jk})_{1\leq j,k\leq m}$ with entries

$$A_{jk} = \begin{cases} 1 & \text{if } P(k) \subset \overline{T(P(j))}, \\ 0 & \text{otherwise.} \end{cases}$$

Denote by $\Sigma_A^{\mathbb{N}} \subset \{1, 2, \ldots, m\}^{\mathbb{N}}$ the corresponding symbolic space induced by A and Σ_A^n the set of words of length n in $\Sigma_A^{\mathbb{N}}$. The projection π from $\Sigma_A^{\mathbb{N}}$ to Λ is given by

$$\omega = (\omega_n)_{n \geq 0} \quad \mapsto \quad \pi(\omega) = \bigcap_{n=0}^{\infty} f^{-n}(P(\omega_n))\,.$$

In view of (57), it follows that any given word of length n gives rise to at least $\left[\frac{\beta(T)}{2}\right] - 2$ words of length $(n+1)$. Hence, we have that

$$\#\Sigma_A^n \geq m \left(\frac{\beta(T)}{2} - 3 \right)^{n-1},$$

which implies that the topological entropy $h_{\text{top}}(T|_\Lambda)$ of $T|_\Lambda$ is at least $\log\left(\frac{\beta(T)}{2} - 3\right)$. This together with Bowen's definition of topological entropy (see [5], [16, page 230]) and the fact that the absolute value of the slope of $T|_\Lambda$ is a constant (namely $\beta(T) > 8$), implies that

$$\dim_{\mathrm{H}} \Lambda = \frac{h_{\text{top}}(T|_\Lambda)}{\log \beta(T)} \geq \frac{\log\left(\frac{\beta(T)}{2} - 3\right)}{\log \beta(T)} \geq 1 - \frac{\log 8}{\log \beta(T)}\,. \quad \square$$

The following result provides a lower bound for the Hausdorff dimension of shrinking target sets associated with piecewise linear maps.

Proposition 8. *Let T be a piecewise linear map on $[0,1]$ and assume that $\beta(T) > 8$. Let $(\Lambda, T|_\Lambda, \mathcal{P}_\Lambda)$ be the associated Markov subsystem arising from Proposition 7. Suppose there exists a compact set $K \supseteq \Lambda$ and an integer $k_0 > 0$, so that $T^{k_0}(P) \supseteq K$ for any interval $P \in \mathcal{P}_\Lambda$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive function and $a \in K$. Then*

$$\dim_{\mathrm{H}} W(T, \psi, a) \geq \frac{1 - \log 8/\log \beta(T)}{1 + \lambda/\log \beta(T)},$$

where $\lambda = \lambda(\psi)$ is the lower order at infinity of the function ψ and

$$W(T, \psi, a) := \{x \in [0,1] : |T^n x - a| \leq \psi(n) \ \text{ for infinitely many } n \in \mathbb{N}\}.$$

Proof. We are given that $(\Lambda, T|_\Lambda, \mathcal{P}_\Lambda)$ is a Markov subsystem of the dynamical system $([0,1], T)$ coming from Proposition 7. Indeed, $\mathcal{P}_\Lambda = \{P(i) \ : \ 1 \leq i \leq m\}$ where each $P(i)$ is an interval and $T|_{P(i)}$ is linear. As in the proof of Proposition 7, denote by $\Sigma_A^{\mathbb{N}} \subset \{1, 2, \ldots, m\}^{\mathbb{N}}$ the corresponding symbolic space of the dynamics of $T|_\Lambda$ induced by the transition matrix A and Σ_A^n the set of words of length n in $\Sigma_A^{\mathbb{N}}$. With this in mind, given a word $(i_0 i_1 \cdots i_{n-1}) \in \Sigma_A^n$, let

$$P(i_0 i_1 \cdots i_{n-1}) := P(i_0) \cap T^{-1}\big(P(i_1)\big) \cap \cdots \cap T^{-(n-1)}\big(P(i_{n-1})\big)$$

and for each $n \in \mathbb{N}$, let

$$\mathcal{P}_n := \Big\{P(i_0 i_1 \cdots i_{n-1}) : (i_0 i_1 \cdots i_{n-1}) \in \Sigma_A^n\Big\}$$

denote the collection of cylinder sets of length n. Now by the Markov property of $\mathcal{P}_\Lambda$, for any cylinder $P(i_0 i_1 \cdots i_{n-1}) \in \mathcal{P}_n$

$$T^{n-1}\big(P(i_0 i_1 \cdots i_{n-1})\big) = P(i_{n-1}). \tag{58}$$

Then with K and k_0 as in the statement of the proposition, we have that $T^{n-1+k_0}\big(P(i_0 i_1 \cdots i_{n-1})\big) \supseteq K$. It therefore follows that for any $a \in K$, there exists a point $x_{i_0 i_1 \cdots i_{n-1}} \in P(i_0 i_1 \cdots i_{n-1})$ such that $T^{n-1+k_0}(x_{i_0 i_1 \cdots i_{n-1}}) = a$. That is, we can find a preimage of the point a under T^{n+k_0-1} on every cylinder of order n. So for any point $x \in B\Big(x_{i_0 i_1 \cdots i_{n-1}}, \frac{\psi(n+k_0-1)}{\beta(T)^{n+k_0-1}}\Big)$, we have that

$$\begin{aligned}|T^{n+k_0-1}(x) - a| &= |T^{n+k_0-1}(x) - T^{n+k_0-1}(x_{i_0 i_1 \cdots i_{n-1}})| \\ &= \beta(T)^{n+k_0-1}|x - x_{i_0 i_1 \cdots i_{n-1}}| \\ &< \psi(n + k_0 - 1).\end{aligned}$$

Therefore,

$$\limsup_{n\to\infty} \bigcup_{i_0 i_1 \cdots i_{n-1} \in \Sigma_A^n} B\Big(x_{i_0 i_1 \cdots i_{n-1}}, \frac{\psi(n+k_0-1)}{\beta(T)^{n+k_0-1}}\Big) \subset W(T, \psi, a). \tag{59}$$

On the other hand, by (58) we have that

$$\kappa_* \beta(T)^{-(n-1)} \leq |P(i_0 i_1 \cdots i_{n-1})| \leq \kappa^* \beta(T)^{-(n-1)},$$

where $\kappa_* = \min_{1\leq i\leq m} |P(i)|$ and $\kappa^* = \max_{1\leq i\leq m} |P(i)|$. Hence

$$\bigcup_{i_0 i_1 \cdots i_{n-1} \in \Sigma_A^n} B\big(x_{i_0 i_1 \cdots i_{n-1}},\ \kappa^* \beta(T)^{-(n-1)}\big) \supseteq \Lambda \tag{60}$$

and so

$$\limsup_{n\to\infty} \bigcup_{i_0 i_1 \cdots i_{n-1} \in \Sigma_A^n} B\big(x_{i_0 i_1 \cdots i_{n-1}},\ \kappa^* \beta(T)^{-(n-1)}\big) \supset \Lambda\,. \tag{61}$$

Now let $\delta := \dim_{\mathrm{H}} \Lambda$ and note that

$$\left(\frac{\psi(n+k_0-1)}{\beta(T)^{n+k_0-1}}\right)^{\frac{s}{\delta}} \geq \kappa^* \beta(T)^{-(n-1)}$$

for any

$$0 < s < s_0 := \limsup_{n\to\infty} \frac{\delta(\log \kappa^* - (n-1)\log \beta(T))}{\log \psi(n+k_0-1) - (n+k_0-1)\log \beta(T)} = \frac{\delta}{1+\lambda/\log \beta(T)} .$$

In other words, for $s < s_0$ the radii of the 's-scaled up' balls associated with (59) are at least the size of the corresponding balls appearing in (61). It then follows via (59), (61) and Proposition 6, that on applying the Mass Transference Principle (the original Theorem 10) with $\mu = \mathcal{H}^\delta|_\Lambda$, we have that

$$\mathcal{H}^s\big(W(T,\psi,a)\big) = \mathcal{H}^s(\Lambda) = \infty . \tag{62}$$

The right hand most equality is valid since $s < \delta$. Now (62) is true for any $s < s_0$ and so together with Proposition 7 it follows that

$$\dim_{\mathrm{H}} W(T,\psi,a) \geq s_0 = \frac{\dim_{\mathrm{H}} \Lambda}{1+\lambda/\log \beta(T)} \geq \frac{1-\log 8/\log \beta(T)}{1+\lambda/\log \beta(T)}. \quad \square$$

As we shall soon see, Proposition 8 will be instrumental in the proof of Theorem 7. Before moving onto the latter, we establish a technical lemma.

Lemma 9. *Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive decreasing function and let $\lambda = \lambda(\psi)$ be its lower order at infinity. Then, for any positive integer k we have that*

$$\liminf_{n\to\infty} \frac{-\log \psi(kn)}{kn} = \lambda.$$

Proof. Recall, $\lambda := \liminf_{n\to\infty} \frac{-\log \psi(n)}{n}$ and thus there exist infinitely many indices $n \in \mathbb{N}$ such that

$$\psi(n) > \exp\left(-(\lambda+\varepsilon)n\right). \tag{63}$$

Now fix a positive integer $k \geq 2$ and let

$$\xi := \liminf_{n\to\infty} \frac{-\log \psi(kn)}{kn}.$$

Thus, for any $\varepsilon > 0$ there exists an $N_\varepsilon > 0$ such that for every $n > N_\varepsilon$

$$\psi(kn) < \exp\left(-(\xi-\varepsilon)kn\right). \tag{64}$$

By definition, we trivially have that $\xi \geq \lambda$. We claim that if ψ is decreasing then we must have equality. With this in mind, assume on the contrary that $\xi > \lambda$ and set $\varepsilon := \frac{\xi-\lambda}{4}$. For our fixed $k \geq 2$, any arbitrarily positive integer can be written in the form $kn + r$ with $n \in \mathbb{N}$ and $r \in \mathbb{N}$ satisfying $0 \leq r \leq k-1$. By (63), there is an increasing sequence $(kn_i + r_i)_{i\geq 1}$ with $n_i \in \mathbb{N}$ and $0 \leq r_i \leq k-1$ such that

$$\psi(kn_i + r_i) > \exp\left(-(\lambda+\varepsilon)(kn_i + r_i)\right).$$

On the other hand, for any $n \in \mathbb{N}$ and $0 \leq r \leq k-1$ such that $kn > N_\epsilon$, by the decreasing property of ψ and (64), we have that

$$\psi(kn+r) \leq \psi(kn) < \exp\left(-(\xi-\varepsilon)kn\right) = \exp\left(-(\xi-\varepsilon)(kn+r)\right)\exp\left((\xi-\varepsilon)r\right) .$$

Thus, for all i large enough we have that

$$\exp\left(-(\lambda+\varepsilon)(kn_i + r_i)\right) < \exp\left(-(\xi-\varepsilon)(kn_i + r_i)\right)\exp\left((\xi-\varepsilon)r_i\right),$$

which in turn implies that

$$\exp\left(\frac{\xi-\lambda}{2}(kn_i + r_i)\right) < \exp\left((\xi-\varepsilon)r_i\right). \tag{65}$$

Now note that with k fixed, the right-hand side of (65) is bounded since r_i lies in the range from 0 to $k-1$. However, since $\frac{\xi-\lambda}{2} > 0$, the left-hand side of (65) tends to infinity as i tends to infinity and we obtain a contradiction. The upshot is that we must have $\xi = \lambda$, as claimed. $\square$

Proof of Theorem 7. We prove Theorem 7 by estimating the upper and lower bounds for the Hausdorff dimension of $W(T_\beta, \psi, a)$ separately.

The upper bound for $\dim_{\mathrm{H}} W(T, \psi, a)$ essentially follows the same line of argument as within the proof of Lemma 8 with $i = 1$ and the estimate (46) replaced by (47). In short, for any $n \in \mathbb{N}$ the preimage $T_\beta^{-n}\big(B(a, \psi(n))\big)$ consists of N_n intervals $\{I_{n,j} : 1 \leq j \leq N_n\}$ with lengths bounded by $2\psi(n)|\beta|^{-n}$ and in view of Remark 12, for any $\epsilon > 0$ there exists $N_0 \geq 1$ such that for all $n \geq N_0$

$$N_n \leq |\beta|^{n(1+\epsilon)}.$$

Now for any $N \geq 1$, we have that

$$W(T_\beta, \psi, a) \subset \bigcup_{n=N}^{\infty} \bigcup_{j=1}^{N_n} I_{n,j} .$$

Thus, given $\rho > 0$ and on choosing $N \geq N_0$ sufficiently large so that $|\beta|^{-N} < \rho$, it follows that for any $s > 0$

$$\mathcal{H}^s_\rho(W(T_\beta,\psi,a)) \le \sum_{n=N}^{\infty}\sum_{j=1}^{N_n}|I_{n,j}|^s \le \sum_{n=N}^{\infty}|\beta|^{n(1+\epsilon)}(2\psi(n)|\beta|^{-n})^s$$
$$\le \sum_{n=N}^{\infty}|\beta|^{n(1+\epsilon)-ns+s\frac{\log\psi(n)}{\log|\beta|}}.$$

Hence, for any $s > \frac{(1+\epsilon)\log|\beta|}{\lambda+\log|\beta|}$ we have that $\mathcal{H}^s(W(T_\beta,\psi,a)) = 0$ and thus

$$\dim_{\mathrm{H}} W(T_\beta,\psi,a) \le \frac{(1+\epsilon)\log|\beta|}{\lambda+\log|\beta|}.$$

Since $\epsilon > 0$ is arbitrary, we obtain the desired upper bound for the dimension of $W(T_\beta,\psi,a)$.

To prove the complementary lower bound, we make use of Proposition 8. With this in mind, for any real number β with $|\beta| > 1$, the transformation T_β can be considered as a piecewise linear mapping of the unit interval $[0,1]$ with $\beta(T) = |\beta|$. Strictly, speaking T_β is defined on $[0,1)$ but we can naturally include the end point one by defining $T_\beta(1) = \beta \pmod 1$. This extension will not effect the dimension of $W(T_\beta,\psi,a)$ since it introduces at most a single point. Now choose $k \in \mathbb{N}$ large enough so that $|\beta|^k > 8$, and note that

$$W(T_\beta,\psi,a) \supseteq W(T^k_\beta,\varphi_k,a) \quad \text{where} \quad \varphi_k(n) := \psi(kn).$$

Let $(\Lambda_k, T^k_\beta|_{\Lambda_k}, \mathcal{P}_{\Lambda_k})$ be the Markov subsystem of $([0,1], T^k_\beta)$ arising from Proposition 7. The following claim will enable us to establish the hypotheses within Proposition 8 regarding the existence of a compact set $K \supseteq \Lambda_k$ and an integer $k_0 > 0$, so that $T^{k_0}_\beta(P) \supseteq K$ for any interval $P \in \mathcal{P}_{\Lambda_k}$. As usual, we let $K(\beta)$ denote the support of the Parry-Yrrap measure μ_β. Recall, that $K(\beta)$ is either the unit interval or a finite union of closed intervals – see Proposition 3 in §3.3.

Claim. *For any interval $I \subseteq [0,1]$, there exists an integer $k(I) > 0$, such that $T^{k(I)}_\beta(I) \supseteq K(\beta)$.*

Proof of Claim. We will use the fact that for any $|\beta| > 1$, the map T_β is locally eventually onto (or topologically exact); i.e. for every non-degenerate subinterval $I \subseteq K(\beta)$ there exists a non-negative integer k such that $T^k(I) \supseteq K(\beta)$. For $\beta > 1$, this is explicitly stated and proved in the work of Troubetzkoy & Varandas [46, Section 3.3] and since $K(\beta) = [0,1]$ when $\beta > 1$ it directly establishes the claim. On the other hand, for $\beta < -1$ the fact is explicitly stated and proved in the work of Liao & Steiner [33, Theorem 2.2]. As a consequence, given any interval $I \subseteq [0,1]$, if $I \cap K(\beta)$ contains an interval then we are done. So assume that this is not the situation. Then, $I \cap \big([0,1] \setminus K(\beta)\big)$ contains an interval and to continue we consider two situations:

(a) There exists a positive integer $\ell(I)$ such that $T_\beta^{\ell(I)}(I) \cap K(\beta)$ contains an interval. In this case the claim follows directly from the locally eventually onto property of T_β.

(b) If (a) does not hold, then for all $n \in \mathbb{N}$, $T^n(I)$ is contained in $[0,1] \setminus K(\beta)$ except for a finite number of points. Therefore

$$\lim_{n\to\infty} m_1\big(T^{-n}([0,1] \setminus K(\beta))\big) \geq m_1(I) > 0\,,$$

where as usual m_1 is one-dimensional Lebesgue measure. However, this contradicts the second assertion of [33, Theorem 2.2]; namely that $\lim\limits_{n\to\infty} m_1(T^{-n}([0,1]\setminus K(\beta))) = 0$. $\square$

On using the above claim, it follows that for any interval $P(i) \in \mathcal{P}_{\Lambda_k} := \{P(i) : 1 \leq i \leq m\}$ there exists an integer $k_0(i) > 0$ such that $T_\beta^{k_0(i)}\big(P(i)\big) \supseteq K(\beta)$. The upshot of this is that the hypotheses within Proposition 8 is satisfied with $K = K(\beta)$ and $k_0 = \max\limits_{1\leq i\leq m} k_0(i)$. Then on applying Proposition 8, we have that

$$\dim_{\mathrm{H}} W(T_\beta, \psi, a) \;\geq\; \dim_{\mathrm{H}} W(T_\beta^k, \varphi_k, a) \;\geq\; \frac{1 - \log 8/(k \log \beta(T))}{1 + \lambda_k/(k \log \beta(T))}$$

where $\lambda_k := \liminf_{n\to\infty} \frac{-\log \varphi_k(n)}{n}$ is the lower order at infinity of φ_k. Now by Lemma 9, since ψ is a real positive decreasing function, we have that $\lambda_k/k = \lambda$ and so on letting $k \to \infty$ we obtain the desired lower bound for the dimension of $W(T_\beta, \psi, a)$. $\square$

5. Final comments

In this section we discuss various natural problems that arise as a consequence of the results proved in this paper. The measure results (namely, Theorems 2 - 5) for matrix transformations are reasonably complete so the problems listed below are essentially concerned with Hausdorff dimension.

5.1. Dimension problem for property (**P**) *targets sets*

Theorem 12 and Theorem 9 give the Hausdorff dimension of shrinking target set $W(T, \{E_n\})$ when the targets sets $\{E_n\}_{n\in\mathbb{N}}$ are a sequence of rectangles or hyperboloids. It is easily seen that both these "shapes" when centred at the origin satisfy the property (**P**) condition of Gallagher [20] adapted for the torus: a subset E of $\mathbb{T}^d$ is said to have property (**P**) if whenever $\mathbf{x} = (x_1, \ldots, x_d) \in E$ and $\|x_i'\| \leq x_i$ $(1 \leq i \leq d)$ then $\mathbf{x}' = (x_1', \ldots, x_d') \in E$. Geometrically, the property simply means that the rectangle $B(0, x_1) \times \ldots \times B(0, x_d)$ is contained within E. In short, it would be desirable to extend and thereby unify our dimension results (with $\mathbf{a} := (0, \ldots, 0)$ in the first instance) to target sets satisfying property (**P**). We now briefly describe what we have in mind.

Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal and all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ are strictly larger than 1. Assume that $1 < \beta_1 \leq \beta_2 \leq \cdots \leq \beta_d$. Let $\mathcal{P}$ be any collection of subsets E of $\mathbb{T}^d$ satisfying property (**P**). Then, for any sequence $\{E_n\}_{n\in\mathbb{N}}$ in $\mathcal{P}$, Theorem 4 implies that $m_d\big(W(T, \{E_n\})\big) = 0$ if $\sum_{n=1}^{\infty} m_d\big(E_n\big) < \infty$. Thus, whenever the measure sum converges, it is natural to ask for the Hausdorff dimension of $W(T, \{E_n\})$. Given that property (**P**) is intimately tied up with rectangles, it is not unreasonable to expect that $\dim_{\mathrm{H}} W(T, \{E_n\})$ is in someway related to the Hausdorff dimension of the 'rectangular' shrinking targets sets $W(T, \Psi, \mathbf{a})$ given by Theorem 12. With this in mind, for any sequence $\{E_n\}_{n\in\mathbb{N}}$ in $\mathcal{P}$, we propose the following candidate for the dimension formula:

$$\dim_{\mathrm{H}} W(T, \{E_n\}) \;=\; \sup_{\substack{\Psi\,:\,\forall\, n\in\mathbb{N} \\ R(0,\Psi(n))\subseteq E_n}} \dim_{\mathrm{H}} W(T, \Psi, 0). \tag{66}$$

Here, as in §4.2, given $\Psi := (\psi_1, \ldots, \psi_d)$ and some point $\mathbf{a} := (a_1, \ldots, a_d) \in \mathbb{T}^d$, for $n \in \mathbb{N}$ we let

$$R\big(\mathbf{a}, \Psi(n)\big) := \Big\{\mathbf{x} \in \mathbb{T}^d : \|x_i - a_i\| \leq \psi_i(n)\ (1 \leq i \leq d)\Big\}.$$

Observe, that since in (66) the supremum is over Ψ such that the corresponding rectangles $R\big(0, \Psi(n)\big)$ are a subset of the sets E_n satisfying property (**P**), we automatically obtain the desired lower bound statement:

$$\dim_{\mathrm{H}} W(T, \{E_n\}) \;\geq\; \sup_{\substack{\Psi\,:\,\forall\, n\in\mathbb{N} \\ R(0,\Psi(n))\subseteq E_n}} \dim_{\mathrm{H}} W(T, \Psi, 0). \tag{67}$$

Thus, establishing (66) boils down to establishing the complimentary upper bound statement.

It is not difficult to see that the dimension formula (66) holds when the targets sets $\{E_n\}_{n\in\mathbb{N}}$ are a sequence of rectangles as in Theorem 12 or hyperboloids as in Theorem 9. The former is obvious. Regarding the latter, for $n \in \mathbb{N}$ we let $\Psi(n) := (1, \cdots, 1, \psi(n))$. Then,

$$R\big(0, \Psi(n)\big) = B(0, 1) \times \cdots \times B(0, 1) \times B(0, \psi(n)) \tag{68}$$

and with reference to Theorem 12

$$\mathcal{U}(\Psi) = \left\{(0, 0, \cdots, 0, t_d) : t_d \text{ is an accumulation point of } \left\{-\frac{\log \psi(n)}{n}\right\}_{n\geq 1}\right\}.$$

Furthermore, for $1 \leq i \leq d-1$, we have that $\mathcal{K}_1(i) = \{i+1, i+2, \cdots, d\}$, $\mathcal{K}_2(i) = \{1, 2, \cdots, i\}$, and $\mathcal{K}_3(i) = \emptyset$. Hence,

$$\theta_1(0, 0, \cdots, 0, t_d) \;=\; \cdots \;=\; \theta_{d-1}(0, 0, \cdots, 0, t_d) \;=\; d\,.$$

For $i = d$, we have that $\mathcal{K}_1(d) = \mathcal{K}_3(d) = \emptyset$ and $\mathcal{K}_2(d) = \{1, 2, \cdots, d\}$. Thus,

$$\theta_d(0, 0, \cdots, 0, t_d) \;=\; d - 1 + \frac{\log \beta_d}{t_d + \log \beta_d}.$$

Therefore, on applying Theorem 12 we obtain that

$$\begin{aligned}\dim_{\mathrm{H}} W(T, \Psi, 0) &= \sup_{t_d} \min_{1 \le i \le d} \theta_i(0, 0, \cdots, 0, t_d) \;=\; \sup_{t_d} \left\{ d - 1 + \frac{\log \beta_d}{t_d + \log \beta_d} \right\} \\ &= d - 1 + \frac{\log \beta_d}{\lambda + \log \beta_d},\end{aligned}$$

where λ is the lower order at infinity of ψ. Since this dimension formula coincides with the dimension of $W^{\times}(T, \Psi, 0)$ given by Theorem 9, and we always have the lower bound (67), we conclude that the supremum in (66) is attained by the choice of rectangles given by (68). In other words, the dimension formula (66) holds when the targets sets $\{E_n\}_{n\in\mathbb{N}}$ are a sequence of hyperboloids as in Theorem 9.

The following is an extension of Gallagher's property $(\mathbf{P})$ condition that naturally incorporates "shapes" not necessarily centred at the origin. Given $\mathbf{a} \in \mathbb{T}^d$, a subset E of $\mathbb{T}^d$ is said to have property $(\mathbf{P_a})$ if whenever $\mathbf{x} = (x_1, \ldots, x_d) \in E$ and $\|x_i' - a_i\| \le x_i$ $(1 \le i \le d)$ then $\mathbf{x}' = (x_1', \ldots, x_d') \in E$. Now with this in mind, let $\mathcal{P}_{\mathbf{a}}$ be any collection of subsets E of $\mathbb{T}^d$ satisfying property $(\mathbf{P_a})$. Then, for any sequence $\{E_n\}_{n\in\mathbb{N}} \in \mathcal{P}_{\mathbf{a}}$, we propose that (66) holds with the origin replaced by $\mathbf{a}$. Clearly, such a statement would unify in full our dimension results for rectangular and hyperboloid target sets; that is, not just for when $\mathbf{a} := (0, \ldots, 0)$.

5.2. Dimension problem for diagonal matrices with negative entries

In the one dimensional case, Theorem 7 extends Theorem 6 by incorporating negative eigenvalues. Naturally, it would be desirable to obtain the higher dimensional analogue of Theorem 7. Indeed, this would clearly follow if we could extend Theorem 12 (the "rectangular" generalization Theorem 6) to the situation that all eigenvalues of T are of modulus strictly larger than 1. Formally, we would expect the following statement to hold in which the conditions on the eigenvalues in Theorem 12 are replaced by conditions on the modulus of the eigenvalues.

Claim 1. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal and all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ are of modulus strictly larger than 1. Assume that $1 < |\beta_1| \le |\beta_2| \le \cdots \le |\beta_d|$. For $1 \le i \le d$, let $\psi_i : \mathbb{R}^+ \to \mathbb{R}^+$ be a real positive decreasing function and $\mathbf{a} \in K = \prod_{i=1}^d K(\beta_i)$. Assume that the set $\mathcal{U}(\Psi)$ of accumulation points $\mathbf{t} = (t_1, t_2, \ldots, t_d)$ of the sequence $\left\{(-\frac{\log \psi_1(n)}{n}, \cdots, -\frac{\log \psi_d(n)}{n})\right\}_{n\ge1}$ is bounded. Then*

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) = \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\le i\le d} \{\theta_i(\mathbf{t})\},$$

where

$$\theta_i(\mathbf{t}) := \sum_{k\in\mathcal{K}_1(i)} 1 + \sum_{k\in\mathcal{K}_2(i)} \left(1 - \frac{t_k}{\log|\beta_i| + t_i}\right) + \sum_{k\in\mathcal{K}_3(i)} \frac{\log|\beta_k|}{\log|\beta_i| + t_i}$$

and, in turn

$$\mathcal{K}_1(i) := \{1\le k\le d : \log|\beta_k| > \log|\beta_i| + t_i\},$$
$$\mathcal{K}_2(i) := \{1\le k\le d : \log|\beta_k| + t_k \le \log|\beta_i| + t_i\},$$

and

$$\mathcal{K}_3(i) := \{1,\dots,d\} \setminus (\mathcal{K}_1(i)\cup\mathcal{K}_2(i)).$$

The key problem with allowing negative eigenvalues is that we do not have an analogue of Fact BW in Section 4.2.1 for negative β-transformations. This fact played a key role our proofs of the upper bound (Proposition 4) and lower bound (Proposition 5) statements for the Hausdorff dimension of $\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a})$. However, by exploiting the framework of Markov subsystems used in proving Theorem 7, it is not too difficult to establish the lower bound of the above claim; that is to say, we can bypass Fact BW altogether and prove that

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \ge \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\le i\le d} \{\theta_i(\mathbf{t})\}.$$

Indeed, for each $1\le i\le d$, by Proposition 7 there exists a Markov subsystem $(\Lambda^{(i)}, T_{\beta_i}|_{\Lambda^{(i)}}, \mathcal{P}_{\Lambda^{(i)}})$ of $([0,1), T_{\beta_i})$ under the assumption that $|\beta_i| > 8$. Also, in view of Proposition 6 we know that the measure $\mathcal{H}^{\delta_i}|_{\Lambda^{(i)}}$ is δ_i-Ahlfors regular where $\delta_i = \dim_{\mathrm{H}} \Lambda^{(i)}$. Now, let $S_i = T_{\beta_i}|_{\Lambda^{(i)}}$ and consider restricted shrinking target set

$$W^*(T,\Psi,\mathbf{a}) := \left\{\mathbf{x}\in\prod_{i=1}^{d}\Lambda^{(i)} : |S_i^n x_i - a_i| \le \psi_i(n)\ (1\le i\le d)\ \text{ for infinitely many } n\in\mathbb{N}\right\}.$$

Then, by definition,

$$W^*(T,\Psi,\mathbf{a}) \subset W(T,\Psi,\mathbf{a}).$$

The first goal is obtain a lower bound for $\dim_{\mathrm{H}} W^*(T,\Psi,\mathbf{a})$. For this, we follow the basic strategy used in proving Proposition 5. However, the key in executing the strategy lies in the fact that each map S_i $(1\le i\le d)$ satisfies the hypothesis of Proposition 8 – this follows on using the same arguments used at the end of the proof of Theorem 7 to show that the Markov subsystem $(\Lambda_m, T_\beta^m|_{\Lambda_m}, \mathcal{P}_{\Lambda_m})$ of $([0,1], T_\beta^m)$ arising from Proposition 7 satisfies the hypotheses of Proposition 8. Then, on naturally adapting the arguments leading to (59) within the proof of Proposition 8, it follows that for each $1 \le i \le d$ there exists an integer $k_0^{(i)}$ such that

$$W^*(T,\Psi,\mathbf{a}) \supseteq \limsup_{n\to\infty} \bigcup_{j_1=1}^{M_{1,n}} \cdots \bigcup_{j_d=1}^{M_{d,n}} \prod_{i=1}^{d} B\Big(x_{n,j_i}^{(i)}, \frac{\psi_i(n-1+k_0^{(i)})}{|\beta_i|^{n-1+k_0^{(i)}}}\Big), \tag{69}$$

where $\{x_{n,j_i}^{(i)}, 1 \le j_i \le M_{i,n}\}$ are the preimages of a_i under S_i^n that fall within cylinders of order n for S_i and $M_{i,n}$ is the number of such cylinders. This is the analogue of the inclusion (53) in the proof of Proposition 5. Now in view of (60) within the proof of Proposition 8, it follows that for each $1 \le i \le d$ there exists a constant κ_i^* such that

$$\bigcup_{j_i=1}^{M_{i,n}} B\big(x_{n,j_i}^{(i)}, \kappa_i^* |\beta_i|^{-n}\big) \supseteq \Lambda^{(i)}\,.$$

In particular, this leads to the following analogue of (56) in the proof of Proposition 5, for any fixed $\mathbf{t} = (t_1,\ldots,t_d) \in \mathcal{U}(\Psi)$:

$$\prod_{i=1}^{d} \Lambda^{(i)} \ \subseteq\ \limsup_{n\to\infty} \bigcup_{j_1=1}^{M_{1,n}} \cdots \bigcup_{j_d=1}^{M_{d,n}} \prod_{i=1}^{d} B\Big(x_{n,j_i}^{(i)}, \Big(\frac{\psi_i(n-1+k_0^{(i)})}{|\beta_i|^{n-1+k_0^{(i)}}}\Big)^{s_i}\ \Big)$$

where

$$0 < s_i < s_0(i) := \frac{\delta_i}{1 + t_i/\log|\beta_i|}\,.$$

The upshot is that given the lim sup set of rectangles appearing on the right hand side of (69), the corresponding lim sup set of '$(s_1,\ldots,s_d)$-scaled up' rectangles satisfies (41) with $p = d, X_i = \Lambda^{(i)}, \delta_i = \dim_{\mathrm{H}} \Lambda^{(i)}$ and $\mu_i = \mathcal{H}^{\delta_i}|_{\Lambda^{(i)}}$ for each $1 \le i \le d$. Thus on applying Theorem 11 with $u_i = (1-\varepsilon)\log|\beta_i|$ and $v_i = t_i$ $(1 \le i \le d)$, we find as in the proof of Proposition 5, that

$$\dim_{\mathrm{H}} W(T,\Psi,\mathbf{a}) \ \ge\ \dim_{\mathrm{H}} W^*(T,\Psi,\mathbf{a}) \ \ge\ \sup_{\mathbf{t}\in\mathcal{U}(\Psi)} \min_{1\le i\le d} \big\{\widehat{\theta}_i(\mathbf{t})\big\}\,,$$

where

$$\widehat{\theta}_i(\mathbf{t}) := \sum_{k\in\mathcal{K}_1(i)} \delta_k + \sum_{k\in\mathcal{K}_2(i)} \delta_k \left(1 - \frac{t_k}{\log|\beta_i| + t_i}\right) + \sum_{k\in\mathcal{K}_3(i)} \frac{\delta_k \log|\beta_k|}{\log|\beta_i| + t_i}$$

and $\mathcal{K}_1(i)$, $\mathcal{K}_2(i)$, $\mathcal{K}_3(i)$ are defined as in Claim 1.

To obtain the desired lower bound, for each $1 \le i \le d$ we need to (i) overcome the underlying assumption that $|\beta_i| > 8$ in the argument above and (ii) replace $\widehat{\theta}_i(\mathbf{t})$ by $\theta_i(\mathbf{t})$ in the above lower bound estimate; i.e., replace δ_i by 1 in the definition of $\widehat{\theta}_i(\mathbf{t})$. As in the proof of Theorem 7, we deal with (i) by working with a high enough iterate $S_i^m := T_{\beta_i}^m|_{\Lambda_m^{(i)}}$ of the map S_i and replacing $\psi(n)$ by $\varphi_m(n) := \psi(mn)$ and then letting m become arbitrarily large. This also deals with (ii) since by Proposition 7, $\dim_{\mathrm{H}} \Lambda_m^{(i)} \to 1$ as $m \to \infty$.

5.3. *Theorem 12 for unbounded* $\mathcal{U}(\Psi)$

In the statement of Theorem 12, we require that the set $\mathcal{U}(\Psi)$ of accumulation points is bounded. In short, this allows us to directly exploit the 'rectangles to rectangles' Mass Transference Principle (Theorem 11). However, this is a matter of convenience and it should be possible to obtain a general form of Theorem 12 (and indeed Claim 1 in §5.2) without assuming that $\mathcal{U}(\Psi)$ is bounded. Indeed, by adapting the arguments used in this paper we can "directly" establish various partial statements. These we now briefly describe.

For $\mathbf{t} = (t_1, \ldots, t_d) \in (\mathbb{R}^+ \cup \{\infty\})^d$, let

$$\mathcal{L}_1(\mathbf{t}) := \{1 \le i \le d : t_i < +\infty\} \quad \text{and} \quad \mathcal{L}_2(\mathbf{t}) := \{1 \le i \le d : t_i = +\infty\}.$$

In turn, with $\theta_i(\mathbf{t})$ as in the statement of Theorem 12, define $\tilde{\theta}_i(\mathbf{t})$ to be the reduced value of $\theta_i(\mathbf{t})$ obtained by removing the "infinite" directions associated with $\mathcal{L}_2(\mathbf{t})$; i.e.,

$$\tilde{\theta}_i(\mathbf{t}) := \sum_{k \in \mathcal{K}_1(i) \setminus \mathcal{L}_2(\mathbf{t})} 1 + \sum_{k \in \mathcal{K}_2(i) \setminus \mathcal{L}_2(\mathbf{t})} \left(1 - \frac{t_k}{\log|\beta_i| + t_i}\right) + \sum_{k \in \mathcal{K}_3(i) \setminus \mathcal{L}_2(\mathbf{t})} \frac{\log|\beta_k|}{\log|\beta_i| + t_i}$$

Then, under the setting of Theorem 12 *but without the assumption that* $\mathcal{U}(\Psi)$ *is bounded*, we are able to adapt the proofs of Propositions 4 and 5 to show that:

$$\sup_{\mathbf{t} \in \mathcal{U}(\Psi)} \min\Big\{ \min_{i \in \mathcal{L}_1(\mathbf{t})} \{\tilde{\theta}_i(\mathbf{t})\},\ \#\mathcal{L}_1(\mathbf{t})\Big\} \le \dim_{\mathrm{H}} W(T, \Psi, \mathbf{a}) \le \sup_{\mathbf{t} \in \mathcal{U}(\Psi)} \min\Big\{ \min_{i \in \mathcal{L}_1(\mathbf{t})} \{\theta_i(\mathbf{t})\},\ \#\mathcal{L}_1(\mathbf{t})\Big\}. \tag{70}$$

Clearly, in the case $\mathcal{U}(\Psi)$ is bounded the upper and lower estimates in (70) coincide. In the unbounded case, this is not necessarily true and so the estimates do not in general provide a precise formula for the dimension. We illustrate this with a concrete example. Let $d = 2$ and T to be the diagonal matrix with entries $\beta_1 = 2$ and $\beta_2 = 3$. Also, given a real number $t_1 > 0$, let $\psi_1(n) = e^{-nt_1}$ and $\psi_2(n) = e^{-n^2}$. Then, it is easily verified that (70) implies that for any $t_1 > 0$

$$\frac{\log 2}{\log 2 + t_1} \le \dim_{\mathrm{H}} W(T, \Psi, \mathbf{a}) \le \min\left\{1, \frac{\log 2 + \log 3}{\log 2 + t_1}\right\}. \tag{71}$$

To the best of our knowledge the precise formula for the dimension is unknown and is not a consequence of know results in the theory of Diophantine approximation. In a forthcoming paper [31], by using the ‘old school” approach of constructing optimal Cantor-type subsets of the set under consideration and applying the Mass Distribution Principle [14, Section 4.1], it is shown that

$$\dim_{\mathrm{H}} W(T, \Psi, \mathbf{a}) = \min\left\{1, \frac{\log 2 + \log 3}{\log 2 + t_1}\right\};$$

that is, the upper bound in (71) is sharp. In general, we are therefore lead to believe that the upper bound in (70) is sharp. In [31], we show that this is indeed the case.

For the sake of completeness, we mention that in [31] we also address the analogous “unbounded” problem in the classical theory of simultaneous Diophantine approximation. For example, given a real number $\tau > 0$, let $S(\tau)$ denote the set of $(x_1, x_2) \in \mathbb{R}^2$ for which the inequalities

$$\|nx_1\| < n^{-\tau} \qquad \text{and} \qquad \|nx_2\| < e^{-n}$$

hold for infinitely many $n \in \mathbb{N}$. Then it follows from known “classical” statements (see for example [41]) that $\dim S(\tau) = 1$ for $1/2 \leq \tau \leq 1$, and that for $\tau > 1$

$$\frac{2}{1+\tau} \leq \dim S(\tau) \leq \min\left\{1, \frac{3}{1+\tau}\right\}.$$

However, to the best of our knowledge we do not have a precise formula for the dimension when $\tau > 1$. In [31], it is shown that for $\tau \geq 1/2$

$$\dim S(\tau) = \min\left\{1, \frac{3}{1+\tau}\right\}.$$

5.4. Badly approximable sets

Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that all eigenvalues of T are of modulus strictly larger than 1 and let $\mathcal{C}$ be any collection of subsets E of $\mathbb{T}^d$ satisfying the bounded property (**B**). For any sequence $\{E_n\}_{n\in\mathbb{N}}$ of subsets in $\mathcal{C}$, we can consider the *badly approximable set with respect to the sequence* $\{E_n\}_{n\in\mathbb{N}}$ as follows:

$$\mathbf{Bad}\big(T, \{E_n\}\big) := \left\{\mathbf{x} \in \mathbb{T}^d : \exists\, n_0(\mathbf{x}) \in \mathbb{N} \ \text{ such that } \ T^n(\mathbf{x}) \notin E_n \ \ \forall\, n \geq n_0(\mathbf{x})\right\}.$$

It is easily seen that the set $\mathbf{Bad}\big(T, \{E_n\}\big)$ is the complement of shrinking target set $W\big(T, \{E_n\}\big)$ and consists of points $\mathbf{x} \in \mathbb{T}^d$ whose orbit under T eventually avoids the given sequence of subsets E_n in $\mathcal{C}$. Hence, Theorem 2 provides us a criterion on the zero-one d-dimensional Lebesgue measure of $\mathbf{Bad}\big(T, \{E_n\}\big)$. Indeed, in the case T is diagonal and all eigenvalues are strictly larger than 1, it follows via Theorem 4 that if $\sum_{n=1}^{\infty} m_d(E_n) = \infty$, then

$$m_d\Big(\mathbf{Bad}(T, \{E_n\})\Big) = m_d\Big(\mathbb{T}^d \setminus W(T, \{E_n\})\Big) = 0\,. \tag{72}$$

Thus, whenever the measure sum diverges, it is natural to ask for the Hausdorff dimension of $\mathbf{Bad}(T, \{E_n\})$. We suspect that for a large class of subsets in $\mathcal{C}$, such as those satisfying the stronger property $(\mathbf{P})$, the associated badly approximable sets are of full dimension. Indeed, it is plausible that they are winning sets in the sense of Schmidt's framework of (α, β)–games - see [2, §1.7.2] and references within. Note that there are obvious cases for which $\mathbf{Bad}(T, \{E_n\})$ is empty (for example if $E_n = \mathbb{T}^d$ for all $n \in \mathbb{N}$) and these should naturally be excluded.

To give a little background and to motivate a concrete problem, we consider the special case when the sequence $\{E_n\}_{n\in\mathbb{N}}$ corresponds to balls. More precisely, given $\mathbf{a} \in \mathbb{T}^d$ and a decreasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$, let

$$\begin{aligned}\mathbf{Bad}(T, \psi, \mathbf{a}) &:= \mathbb{T}^d \setminus W(T, \psi, \mathbf{a}) \\ &= \Big\{\mathbf{x} \in \mathbb{T}^d : \liminf_{n\to\infty}\ \psi(n)^{-1}\ \|T^n\mathbf{x} - \mathbf{a}\| > 1\Big\}\,.\end{aligned}$$

Also, let

$$\mathbf{Bad}(T, \mathbf{a}) := \Big\{\mathbf{x} \in \mathbb{T}^d : \liminf_{n\to\infty} \|T^n\mathbf{x} - \mathbf{a}\| > 0\Big\}\,.$$

Then, it is easily seen that if $\psi(n) \to 0$ as $n \to \infty$, then

$$\mathbf{Bad}(T, \mathbf{a}) \subset \mathbf{Bad}\big(T, \mathbf{a}, \psi\big)$$

and so if the badly approximable set $\mathbf{Bad}(T, \mathbf{a})$ has full dimension then so does $\mathbf{Bad}(T, \psi, \mathbf{a})$. With this is mind, Dani [13] showed that if T is a non-singular, semi-simple integer matrix and $\mathbf{a} \in \mathbb{Q}^d/\mathbb{Z}^d$, then $\dim_{\mathrm{H}} \mathbf{Bad}(T, \mathbf{a}) = d$. In fact, he showed that $\mathbf{Bad}(T, \mathbf{a})$ is winning. Dani's winning result was later extended by Broderick, Fishman & Kleinbock [6] to any non-singular, integer matrix transformation and $\mathbf{a} \in \mathbb{T}^d$. Regarding non-integer matrix transformations, we have a complete dimension result in dimension one. Indeed, for any $\beta \in (1, 2]$ and $a \in \mathbb{T}$, Färm, Persson & Schmeling [17] have shown that $\mathbf{Bad}(T_\beta, a)$ is "strong" winning and hence has full Hausdorff dimension. Subsequently, Yang & Wang [53] extended the full Hausdorff dimension result to any $\beta > 1$. To the best of our knowledge, the problem of determining $\dim_{\mathrm{H}} \mathbf{Bad}(T, \mathbf{a})$ when T is a real, non-singular matrix transformation of $\mathbb{T}^d$ with $d \geq 2$ is open. In fact, it seems that the dimension result is currently unknown even in the case that T is a diagonal matrix with all eigenvalues strictly larger than 1.

Now let us consider the special case when the sequence $\{E_n\}_{n\in\mathbb{N}}$ corresponds to hyperboloids. For the sake of simplicity, suppose that $T = \mathrm{diag}\,(t_1, \ldots, t_d)$ is an integer, diagonal matrix with $t_i \geq 2$. Then in line with the discussion above for balls, given $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{T}^d$ and a decreasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$, we consider the sets

$$\begin{aligned}\mathbf{Bad}^\times(T, \psi, \mathbf{a}) &:= \mathbb{T}^d \setminus W^\times(T, \psi, \mathbf{a}) \\ &= \Big\{\mathbf{x} \in \mathbb{T}^d : \liminf_{n\to\infty} \ \psi(n)^{-1} \prod_{1\leq i\leq d} \|t_i^n x_i - a_i\| > 1\Big\}\end{aligned}$$

and

$$\mathbf{Bad}^\times(T, \mathbf{a}) := \Big\{\mathbf{x} \in \mathbb{T}^d : \liminf_{n\to\infty} \prod_{1\leq i\leq d} \|t_i^n x_i - a_i\| > 0\Big\}.$$

If $\psi(n) \to 0$ as $n \to \infty$, then $\mathbf{Bad}^\times(T, \mathbf{a}) \subset \mathbf{Bad}^\times\big(T, \mathbf{a}, \psi\big)$ and so the aim is to show that the multiplicative badly approximable set $\mathbf{Bad}^\times(T, \mathbf{a})$ has full dimension. Given the one dimension result for balls (namely that $\dim \mathbf{Bad}(T_\beta, a) = 1$), this is relatively straightforward to establish. Indeed, we start with the observation that

$$\liminf_{n\to\infty} \prod_{1\leq i\leq d} \|t_i^n x_i - a_i\| \geq \prod_{1\leq i\leq d} \liminf_{n\to\infty} \|t_i^n x_i - a_i\|.$$

Thus, it follows that

$$\prod_{1\leq i\leq d} \mathbf{Bad}(t_i, a_i) \subseteq \mathbf{Bad}^\times(T, \mathbf{a}),$$

where for each $1 \leq i \leq d$

$$\mathbf{Bad}(t_i, a_i) := \Big\{x_i \in \mathbb{T} : \liminf_{n\to\infty} \|t_i^n x_i - a_i\| > 0\Big\}.$$

In turn, since each $\mathbf{Bad}(t_i, a_i)$ has Hausdorff dimension 1, we obtain that

$$\dim_{\mathrm{H}} \mathbf{Bad}^\times(T, \mathbf{a}) \geq \dim_{\mathrm{H}} \prod_{1\leq i\leq d} \mathbf{Bad}(t_i, a_i) \geq \sum_{1\leq i\leq d} \dim_{\mathrm{H}} \mathbf{Bad}(t_i, a_i) = d\,.$$

The complementary upper bound statement is trivial. Thus, $\dim_{\mathrm{H}} \mathbf{Bad}^\times(T, \mathbf{a}) = d$ as desired.

We now describe a class of sequences $\{E_n\}_{n\in\mathbb{N}}$ that naturally unify the above badly approximable sets for balls and hyperboloids. At the same time it allows us to state a concrete problem. Suppose that E is a subset of $\mathbb{T}^d$ satisfying the bounded property (**B**) and furthermore suppose that E contains the origin. Next, given $\mathbf{a} \in \mathbb{T}^d$ and a decreasing function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$, for each $n \in \mathbb{N}$ let

$$E_n(\mathbf{a}, \psi) = \mathbf{a} + \psi(n)E := \big\{\mathbf{a} + \psi(n)\mathbf{x} : \mathbf{x} \in E\big\}.$$

Note that if E satisfies Gallagher's property $(\mathbf{P})$ condition then for each $n \in \mathbb{N}$, the set $E_n(\mathbf{a}, \psi)$ satisfies the general property $(\mathbf{P_a})$ introduced in §5.1. Now let

$$\mathbf{Bad}(T, \psi, \mathbf{a}, E) := \left\{\mathbf{x} \in \mathbb{T}^d : \exists\, n_0(\mathbf{x}) \in \mathbb{N} \text{ such that } T^n(\mathbf{x}) \notin E_n(\mathbf{a}, \psi) \;\; \forall\, n \geq n_0(\mathbf{x})\right\}$$

denote the badly approximable set with respect to the sequence $\{E_n(\mathbf{a}, \psi)\}_{n \in \mathbb{N}}$. Furthermore, let

$$\mathbf{Bad}(T, \mathbf{a}, E) := \left\{\mathbf{x} \in \mathbb{T}^d : \exists\, c(\mathbf{x}) > 0 \text{ such that } T^n(\mathbf{x}) \notin \mathbf{a} + c(\mathbf{x})E \;\; \forall\, n \in \mathbb{N}\right\}.$$

It is easily verified, that if we take E to be the ball $B(\mathbf{0}, 1)$ (resp. the hyperbola $H(\mathbf{0}, 1)$) then $\mathbf{Bad}(T, \psi, \mathbf{a}, E)$ coincides with $\mathbf{Bad}(T, \psi, \mathbf{a})$ (resp. $\mathbf{Bad}^\times(T, \psi, \mathbf{a})$) and $\mathbf{Bad}(T, \mathbf{a}, E)$ coincides with $\mathbf{Bad}(T, \mathbf{a})$ (resp. $\mathbf{Bad}^\times(T, \mathbf{a})$). We suspect that the badly approximable set $\mathbf{Bad}(T, \mathbf{a}, E)$ is of full dimension and thus by default $\mathbf{Bad}(T, \psi, \mathbf{a}, E)$ is also of full dimension. More precisely, we would expect the following statement to hold.

Claim 2. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is integer and all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ are of modulus strictly larger than 1. Then*

$$\dim_{\mathrm{H}} \mathbf{Bad}(T, \mathbf{a}, E) = d\,.$$

It is plausible that the claim is true without the assumption that T is integer. However, as mentioned above, without the integer assumption the problem is currently open even for balls (i.e., when $E = B(\mathbf{0}, 1)$). A potentially interesting starting point towards establishing the claim would be to consider the situation in which E satisfies Gallagher's property $(\mathbf{P})$ condition and T is diagonal with all eigenvalues strictly larger than 1.

We now briefly consider another aspect of the badly approximable theory. Let $c \in (0, 1)$ and with $\mathbf{Bad}(T, \mathbf{a}, E)$ in mind, consider the set

$$\mathbf{Bad}_c(T, \mathbf{a}, E) := \left\{\mathbf{x} \in \mathbb{T}^d : T^n(\mathbf{x}) \notin \mathbf{a} + cE \;\; \forall\, n \in \mathbb{N}\right\}.$$

In short, we fix the so called badly approximable constant $c(\mathbf{x})$ appearing in $\mathbf{Bad}(T, \mathbf{a}, E)$. Then, by definition

$$\mathbf{Bad}(T, \mathbf{a}, E) = \bigcup_{0<c<1} \mathbf{Bad}_c(T, \mathbf{a}, E)\,.$$

When E is an open set, the corresponding set $\mathbf{Bad}_c(T, \mathbf{a}, E)$ is often referred to as a survivor set in the study of (open) dynamical systems. The associated open set $\mathbf{a} + cE$ is referred to as a hole and we are interested in points whose orbit under T avoid the hole. In general, it is difficult to give an exact formula for $\dim_{\mathrm{H}} \mathbf{Bad}_c\big(T, \mathbf{a}, E\big)$ and we are interested in determining how $\dim_{\mathrm{H}} \mathbf{Bad}_c\big(T, \mathbf{a}, E\big)$ varies with respect to the positioning of the hole which is governed by $\mathbf{a}$ and its size which is governed by $0 < c < 1$. So with this in mind, Urbański [49] proved that if T is an expanding map of $\mathbb{T}$ and $E = [0, 1]$ then the dimension function $c \mapsto \dim_{\mathrm{H}} \mathbf{Bad}_c(T, \mathbf{0}, E)$ is a devil's staircase. The same statement was shown to hold by Nilson [36] in the case T is the doubling map, and by Kalla, Kong, Langeveld & Li [27] in the case T is a β-transformation with $\beta \in (1, 2]$. The problem of extending the latter to all $\beta > 1$ and indeed to higher dimensions is clearly a natural path to pursue. In the first instance, establishing a statement of the following type would in our opinion represent serious progress.

Claim 3. *Let T be a real, non-singular matrix transformation of the torus $\mathbb{T}^d$. Suppose that T is diagonal and all eigenvalues $\beta_1, \beta_2, \ldots, \beta_d$ are strictly larger than 1. Furthermore, let E be a subset of $\mathbb{T}^d$ satisfying Gallaghers's property ($\mathbf{P}$) condition. Then the dimension function*

$$c \mapsto \dim_{\mathrm{H}} \mathbf{Bad}_c\big(T, \mathbf{0}, E\big)$$

is a devil's staircase.

Indeed, establishing the claim in the case T is integer and $E = B(\mathbf{0}, 1)$ would be most desirable.

5.5. Shrinking targets restricted to manifolds

For the sake of simplicity, through out this section T will be an integer, non-singular matrix transformation of the torus $\mathbb{T}^d$. Also, we suppose that T is diagonal with eigenvalues $1 < \beta_1 \leq \beta_2 \leq \cdots \leq \beta_d$. Finally, given $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ we consider the "basic" shrinking target set

$$W(T, \psi) = W(T, \psi, 0) := \{\mathbf{x} \in \mathbb{T}^d : T^n(\mathbf{x}) \in B(0, \psi(n)) \ \text{ for infinitely many } n \in \mathbb{N}\}.$$

In view of Theorems 5 and 6, we have a complete description of the "size" of $W(T, \psi)$ in terms of both Lebsegue measure and Hausdorff dimension. Indeed, the former implies that

$$m_d\big(W(T, \psi)\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} \psi(n)^d < \infty \\[2ex] 1 & \text{if } \sum_{n=1}^{\infty} \psi(n)^d = \infty, \end{cases} \tag{73}$$

while the latter implies that

$$\dim_{\mathrm{H}} W(T, \psi) = \min_{1 \leq i \leq d} \theta_i(\lambda). \tag{74}$$

We now add a little twist which is very much in line with the classical theory of Diophantine approximation on manifolds – see [2, Section 6] for background and further references. Suppose that the coordinates of the point $\mathbf{x}$ in $\mathbb{T}^d$ are confined by functional relations or equivalently are restricted to a sub-manifold $\mathcal{M}$ of $\mathbb{T}^d$. We then consider the following two natural problems.

Problem 1. To develop a Lebesgue theory for $\mathcal{M} \cap W(T,\psi)$.

Problem 2. To develop a Hausdorff theory for $\mathcal{M} \cap W(T,\psi)$.

In short, the aim is to establish analogues of (73) and (74) for the set $\mathcal{M} \cap W(T,\psi)$. The fact that the points $\mathbf{x} \in \mathbb{T}^d$ of interest are of dependent variables, which reflects the fact that $\mathbf{x} \in \mathcal{M}$, introduces various difficulties even in the specific case that $\mathcal{M}$ is a planar curve $\mathcal{C}$. However, in this case we have recently obtained a reasonably complete theory. Briefly, assume that $d = 2$ and that the planar curve

$$\mathcal{C} = \mathcal{C}_f := \{(x, f(x)) : x \in [0,1]\}$$

is the graph of a bi-Lipschitz function $f : [0,1] \to \mathbb{R}$. Let m denote the normalised, induced one dimensional Lebesgue measure on $\mathcal{C}$. Then, the main measure result in our forthcoming paper [32] implies that

$$m\big(\mathcal{C} \cap W(T,\psi)\big) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} \psi(n)^d < \infty \\ 1 & \text{if } \sum_{n=1}^{\infty} \psi(n)^d = \infty. \end{cases}$$

As usual, let λ be the lower order at infinity of ψ and recall that $1 < \beta_1 \leq \beta_2$. Then, the main dimension result in [32] implies the following statement. *Assume that* $0 \leq \lambda \leq \log\beta_2$. *Then*

$$\dim\big(\mathcal{C} \cap W(T,\psi)\big) \;\leq\; \frac{1 - \frac{\lambda}{\log\beta_2}}{1 + \frac{\lambda}{\log\beta_2}}\,, \tag{75}$$

and we have equality in (75) *for* $0 \leq \lambda \leq \log\beta_2 - \log\beta_1$. *Moreover, if* $\mathcal{C}$ *is a line with rational slope then we also have equality in* (75) *for* $\log\beta_2 - \log\beta_1 < \lambda \leq \log\beta_2$, *conditional on the validity of the abc-conjecture.*

Remark 13. Let $\mathcal{C}$ be the diagonal line $L := \{(x,x) : x \in [0,1]\}$ and T to be the diagonal matrix with entries $\beta_1 = 2$ and $\beta_2 = 3$. Then, a simple consequence of the above dimension result is the following number theoretic statement which may be of independent interest: for $0 \leq \tau \leq 1$ the set

$$\big\{x \in [0,1] \;:\; \max\{\|2^n x\|, \|3^n x\|\} < 3^{-n\tau} \text{ for infinitely many n} \in \mathbb{N}\big\}$$

has Hausdorff dimension $(1-\tau)/(1+\tau)$. For $\tau > 1-(\log 2/\log 3)$, our proof is conditional on the validity of the *abc*-conjecture.

Two things are worth mentioning. Firstly, for $\log\beta_2 - \log\beta_1 < \lambda \le \log\beta_2$ we suspect that we also have equality in (75) for all "bi-Lipschitz" planar curves (not just rational lines) and almost certainly the use of the *abc*-conjecture is an overkill. Secondly, to the best of our knowledge, beyond the planar case very little seems to be known and in our opinion Problems 1 & 2 represent interesting and potentially fruitful avenues of research.

Acknowledgments

We would like to thank Jérôme Buzzi, Gerhard Keller, Benoît Saussol, and Baowei Wang for the many useful discussions. Also we would like thank Yubin He for reading the paper carefully and pointing out various inconsistencies. Almost certainly there are more errors of one form or another and of course we take full responsibility. SV would like to take this opportunity to thank the wonderful Bridget and the dynamic duo Ayesha and Iona for absolutely everything over the difficult and weird covid years. The duo recently turned twenty-one and their curiosity and optimism remains a marvel – long may it last champions and remember those magical words of Glinda from the land of Oz "You've always had the power my dear, you just had to learn it yourself".

Appendix A. Proof of Theorem 9 by Baowei Wang[3]

We start with stating two lemmas that we will make use of during the course of establishing Theorem 9. As in the main body of the paper, balls are always with respect to the maximum norm and thus correspond to a hypercubes. Indeed, the diameter $d(B)$ of a ball B can equivalently be interpreted as the side length of a hypercube.

Lemma 10. ([4, Lemma 1]) *Let $d \in \mathbb{N}$ and δ be a sufficiently small positive number. Then, for any $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{T}^d$ and $s \in (d-1, d)$ the set*

$$H_d(\mathbf{a}, \delta) = \{\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{T}^d : \|x_1 - a_1\| \cdots \|x_d - a_d\| < \delta\}$$

has a covering $\mathcal{B}$ by d-dimensional balls B such that

$$\sum_{B \in \mathcal{B}} d(B)^s \ll \delta^{s-d+1},$$

where $d(B)$ is the length of a side of U and $\ll$ implies an inequality with a factor independent of δ.

The above lemma does not precisely correspond to the Bovey-Dodson statement [4, Lemma 1]. However, it is readily verified that in establishing Lemma 10 we can, without

[3] School of Mathematics, Huazhong University of Science and Technology, Wuhan 430074, China.
E-mail address: bwei-wang@hust.edu.cn

loss of generality, ignore the 'shift' $\mathbf{a} \in \mathbb{T}^d$. Then, the problem reduces to finding an appropriate cover by balls of the set $\{(x_1, \ldots, x_d) \in [0, 1/2]^d : x_1 \cdots x_d < \delta\}$. In short, for this task the Bovey-Dodson statement is directly applicable.

Lemma 11. ([14, Corollary 7.12]) *Let F be any subset of $\mathbb{R}^d$, and let E be a subset of the x_d-axis. Assume that*

$$\dim_{\mathrm{H}} F \cap L_x \geq t$$

for all $x \in E$, where L_x is the plane parallel to all other axis through the point $(0, \ldots, 0, x)$. Then

$$\dim_{\mathrm{H}} F \geq t + \dim_{\mathrm{H}} E.$$

We now move onto the task of proving Theorem 9. This will be done by establishing the upper and lower bounds for $\dim_{\mathrm{H}} W^{\times}(T, \psi, \mathbf{a})$ separately. Recall, that

$$W^{\times}(T, \psi, \mathbf{a}) := \{\mathbf{x} \in \mathbb{T}^d : T^n(\mathbf{x}) \in H(\mathbf{a}, \psi(n)) \ \text{ for infinitely many } n \in \mathbb{N}\}$$

where $H(\mathbf{a}, \psi(n))$ is the hyperboloid region given by (11).

Proposition 9. *Under the setting of Theorem 9, we have that*

$$\dim_{\mathrm{H}} W^{\times}(T, \Psi, \mathbf{a}) \ \leq \ d - 1 + \frac{\log |\beta_d|}{\lambda + \log |\beta_d|}.$$

Proof. Observe that we can re-write $W^{\times}(T, \Psi, \mathbf{a})$ as

$$W^{\times}(T, \Psi, \mathbf{a}) = \limsup_{n \to \infty} \ E_n^{\times}(T, \psi, \mathbf{a}) \tag{76}$$

where

$$E_n^{\times}(T, \psi, \mathbf{a}) := \{\mathbf{x} \in \mathbb{T}^d : T^n(\mathbf{x}) \in H(\mathbf{a}, \psi(n))\} = \left\{\mathbf{x} \in \mathbb{T}^d : \prod_{i=1}^{d} \|T_{\beta_i}^n x_i - a_i\| < \psi(n)\right\}.$$

As in the main body of the paper, T_{β_i} is the standard β-transformation with $\beta = \beta_i$ and we do not distinguish between β-transformations acting on the unit interval $[0, 1)$ or the torus $\mathbb{T}$. The proof of the proposition relies on finding an "efficient" covering by balls of the lim sup set (76). So with this in mind, for $n \in \mathbb{N}$, we first obtain an efficient cover of the set $E_n^{\times}(T, \psi, \mathbf{a})$.

For any $1 \leq i \leq d$, let $\{C_{n,j}^{(i)} : 1 \leq j \leq N_{i,n}\}$ be the cylinders of order n associated with the transformation T_{β_i}. By definition, these $N_{i,n}$ intervals are disjoint and cover $\mathbb{T}$. Hence

$$\mathbb{T}^d = \bigcup_{j_1=1}^{N_{1,n}} \cdots \bigcup_{j_d=1}^{N_{d,n}} C_{n,j_1}^{(1)} \times \cdots \times C_{n,j_d}^{(d)},$$

where the d-dimensional "rectangles" $C_{n,j_1}^{(1)} \times \cdots \times C_{n,j_d}^{(d)}$ are disjoint. For $n \in \mathbb{N}$, let

$$J_n := \left\{ \mathbf{j} = (j_1, \ldots, j_d) : 1 \le j_i \le N_{i,n} \ (1 \le i \le d) \right\}$$

and for $\mathbf{j} \in J_n$, let

$$E_{n,\mathbf{j}}^{\times}(T, \psi, \mathbf{a}) := \left\{ \mathbf{x} \in C_{n,j_1}^{(1)} \times \cdots \times C_{n,j_d}^{(d)} : \prod_{i=1}^{d} \|T_{\beta_i}^n x_i - a_i\| < \psi(n) \right\}.$$

It follows that

$$E_n^{\times}(T, \psi, \mathbf{a}) = \bigcup_{\mathbf{j} \in J_n} E_{n,\mathbf{j}}^{\times}(T, \psi, \mathbf{a}).$$

By Lemma 10, with $\delta = \psi(n)$ and n sufficiently large, for any $s \in (d-1, d)$ there exists a covering $\mathcal{B}_n$ of the hyperboloid $H\big(\mathbf{a}, \psi(n)\big)$ by balls B such that

$$\sum_{B \in \mathcal{B}_n} d(B)^s \ll \psi(n)^{s-d+1}. \tag{77}$$

By definition

$$E_{n,\mathbf{j}}^{\times}(T, \psi, \mathbf{a}) = \left(T^n|_{C_{n,j_1}^{(1)} \times \cdots \times C_{n,j_d}^{(d)}} \right)^{-1} \left(H\big(\mathbf{a}, \psi(n)\big) \right),$$

and so it follows that

$$E_{n,\mathbf{j}}^{\times}(T, \psi, \mathbf{a}) \subset \left(T^n|_{C_{n,j_1}^{(1)} \times \cdots \times C_{n,j_d}^{(d)}} \right)^{-1} \Big(\bigcup_{B \in \mathcal{B}_n} B \Big) = \bigcup_{B \in \mathcal{B}_n} \left(T^n|_{C_{n,j_1}^{(1)} \times \cdots \times C_{n,j_d}^{(d)}} \right)^{-1} (B).$$

On making use of the fact that for each $1 \le i \le d$, the n-th iteration of T_{β_i} on $C_{n,j_i}^{(i)}$ is an affine function, it can be verified that for any $B \in \mathcal{B}_n$:

$$R_{n,\mathbf{j}}(B) := \left(T^n|_{C_{n,j_1}^{(1)} \times \cdots \times C_{n,j_d}^{(d)}} \right)^{-1} (B)$$

corresponds to either the empty set or to a rectangle with side length $|\beta_i|^{-n} d(B)$ along the x_i-th axis. The upshot is that

$$E_n^{\times}(T, \psi, \mathbf{a}) \subset \bigcup_{\mathbf{j} \in J_n} \bigcup_{B \in \mathcal{B}_n} R_{n,\mathbf{j}}(B),$$

and so for N large enough

$$W^{\times}(T,\psi,\mathbf{a}) \ \subset \ \bigcup_{n=N}^{\infty} E_n^{\times}(T,\psi,\mathbf{a}) \ \subset \ \bigcup_{n=N}^{\infty} \bigcup_{j_1=1}^{N_{1,n}} \cdots \bigcup_{j_d=1}^{N_{d,n}} \bigcup_{B\in\mathcal{B}_n} R_{n,\mathbf{j}}(B)\,.$$

For any $1 \le i \le d$, whenever $R_{n,\mathbf{j}}(B)$ is non-empty, as already mentioned above the side length of the rectangle along the x_i-th axis is $|\beta_i|^{-n}d(B)$ and by assumption $|\beta_i|^{-n}d(B) \ge |\beta_d|^{-n}d(B)$. We now cover the rectangle by balls with diameter equal to the shortest side length of the rectangle. A straightforward geometric argument shows that we can find a collection $\mathcal{C}_n$ of balls with diameter $|\beta_d|^{-n}d(B)$ that cover $R_{n,\mathbf{j}}(B)$ with

$$\#\mathcal{C}_n \ \le \ \prod_{i=1}^{d}\left(\frac{|\beta_i|^{-n}d(B)}{|\beta_d|^{-n}d(B)}+1\right) \ = \ \prod_{i=1}^{d}\left(\frac{|\beta_d|^{n}}{|\beta_i|^{n}}+1\right) \ \le \ 2^d\prod_{i=1}^{d}\frac{|\beta_d|^{n}}{|\beta_i|^{n}}.$$

Thus, given $\rho > 0$ and on choosing N sufficiently large so that $|\beta_d|^{-n}d(B) < \rho$ for all $B \in \mathcal{B}_n$ and for any $n \ge N$, it follows from the definition of s-dimensional Hausdorff measure that for any $s > 0$

$$\begin{aligned}\mathcal{H}^s_\rho(W^{\times}(T,\psi,\mathbf{a})) &\le \sum_{n=N}^{\infty}\sum_{\mathbf{j}\in J_n}\sum_{B\in\mathcal{B}_n}\#\mathcal{C}_n\ \left(|\beta_d|^{-n}d(B)\right)^s \\ &\le \sum_{n=N}^{\infty}\sum_{\mathbf{j}\in J_n}\left(2^d\prod_{i=1}^{d}\frac{|\beta_d|^{n}}{|\beta_i|^{n}}\right)|\beta_d|^{-ns}\sum_{B\in\mathcal{B}_n}d(B)^s. \qquad (78)\end{aligned}$$

Now for any given $\epsilon > 0$, it follows via (47) that for n sufficiently large

$$\#J_n \ \le \ \prod_{i=1}^{d}|\beta_i|^{n(1+\epsilon)}\,.$$

This together with (77), (78) and the assumption that $|\beta_d| \ge |\beta_i| > 1$ for any $1 \le i \le d$, implies that for any $s \in (d-1,d)$ and N sufficiently large

$$\begin{aligned}\mathcal{H}^s_\rho(W^{\times}(T,\psi,\mathbf{a})) &\le \sum_{n=N}^{\infty}\prod_{i=1}^{d}|\beta_i|^{n(1+\epsilon)}\left(2^d\prod_{i=1}^{d}\frac{|\beta_d|^{n}}{|\beta_i|^{n}}\right)|\beta_d|^{-ns}\left(\psi(n)\right)^{s-d+1} \\ &\le 2^d\sum_{n=N}^{\infty}|\beta_d|^{n(d(1+\epsilon)-s)}\left(\psi(n)\right)^{s-d+1} \\ &= 2^d\sum_{n=N}^{\infty}\exp\Big(n\Big(d(1+\epsilon)\log|\beta_d| - (d-1)\frac{\log\psi(n)}{n} \\ &\qquad\qquad -s\Big(\log|\beta_d| - \frac{\log\psi(n)}{n}\Big)\Big)\Big)\,.\end{aligned}$$

Now for any

$$s > d - 1 + \frac{(d\epsilon + 1)\log|\beta_d|}{\lambda + \log|\beta_d|}$$

the above exponential sum converges and so $\mathcal{H}^s(W^\times(T,\psi,\mathbf{a})) = 0$. Therefore, it follows from the definition of Hausdorff dimension that

$$\dim_{\mathrm{H}} W^\times(T,\psi,\mathbf{a}) \le d - 1 + \frac{(d\epsilon + 1)\log|\beta_d|}{\lambda + \log|\beta_d|}.$$

Since $\epsilon > 0$ is arbitrary, on letting $\epsilon \to 0$ we obtain the desired upper bound for $\dim_{\mathrm{H}} W^\times(T,\psi,\mathbf{a})$. $\square$

We now establish the complementary lower bound statement for the Hausdorff dimension of the set $W^\times(T,\psi,\mathbf{a})$.

Proposition 10. *Under the setting of Theorem 9, we have that*

$$\dim_{\mathrm{H}} W^\times(T,\Psi,\mathbf{a}) \ \ge\ d - 1 + \frac{\log|\beta_d|}{\lambda + \log|\beta_d|}\,.$$

Proof. By Theorem 7, for any $a_d \in K(\beta_d)$ we have that

$$\dim_{\mathrm{H}} W^\times(T_{\beta_d},\psi,a_d) = \dim_{\mathrm{H}} W(T_{\beta_d},\psi,a_d) = \frac{\log|\beta_d|}{\lambda + \log|\beta_d|}.$$

Now it is easily verified that for any $x_d \in W^\times(T_{\beta_d},\psi,a_d)$

$$\big([0,1)^{d-1} \times W^\times(T_{\beta_d},\psi,a_d)\big) \cap L_{x_d} = [0,1)^{d-1}\,.$$

Hence, it follows that

$$\dim_{\mathrm{H}} \Big(\big([0,1)^{d-1} \times W^\times(T_{\beta_d},\psi,a_d)\big) \cap L_{x_d}\Big) \ge d - 1.$$

Applying Lemma 11, we obtain

$$\dim_{\mathrm{H}} \Big([0,1)^{d-1} \times W^\times(T_{\beta_d},\psi,a_d)\Big) \ge d - 1 + \frac{\log|\beta_d|}{\lambda + \log|\beta_d|}.$$

This together with the fact that

$$[0,1)^{d-1} \times W^\times(T_{\beta_d},\psi,a_d) \ \subset\ W^\times(T,\psi,\mathbf{a}),$$

implies that

$$\dim_{\mathrm{H}} W^\times(T,\psi,\mathbf{a}) \ge d - 1 + \frac{\log|\beta_d|}{\lambda + \log|\beta_d|}\,. \quad \square$$

References

[1] V. Baladi, Positive Transfer Operators and Decay of Correlations, Advanced Series in Nonlinear Dynamics, vol. 16, World Scientific Publishing Co., Inc., River Edge, NJ, 2000.
[2] V. Beresnevich, F. Ramírez, S. Velani, Metric Diophantine approximation: aspects of recent work, in: D. Badziahin, A. Gorodnik, N. Peyerimhoff (Eds.), Dynamics and Analytic Number Theory, in: LMS Lecture Note Series, vol. 437, Cambridge University Press, 2016, pp. 1–95.
[3] V. Beresnevich, S. Velani, A mass transference principle and the Duffin-Schaeffer conjecture for Hausdorff measures, Ann. Math. (2) 164 (3) (2006) 971–992.
[4] J.D. Bovey, M.M. Dodson, The fractional dimension of sets whose simultaneous rational approximations have errors with a small product, Bull. Lond. Math. Soc. 10 (2) (1978) 213–218.
[5] R. Bowen, Topological entropy for noncompact sets, Trans. Am. Math. Soc. 184 (1973) 125–136.
[6] R. Broderick, L. Fishman, D. Kleinbock, Schmidt's game, fractals, and orbits of toral endomorphisms, Ergod. Theory Dyn. Syst. 31 (4) (2011) 1095–1107.
[7] R.C. Bradley, Basic properties of strong mixing conditions. A survey and some open questions, Probab. Surv. 2 (2005) 107–144.
[8] Y. Bugeaud, B.-W. Wang, Distribution of full cylinders and the Diophantine properties of the orbits in β-expansions, J. Fractal Geom. 1 (2) (2014) 221–241.
[9] J. Buzzi, Intrinsic ergodicity of affine maps in $[0,1]^d$, Monatshefte Math. 124 (2) (1997) 97–118.
[10] J. Buzzi, Absolutely continuous invariant probability measures for arbitrary expanding piecewise $\mathbb{R}$-analytic mappings of the plane, Ergod. Theory Dyn. Syst. 20 (2000) 697–708.
[11] J. Buzzi, Thermodynamical formalism for piecewise invertible maps: absolutely continuous invariant measures as equilibrium states, in: Smooth Ergodic Theory and Its Applications, Seattle, WA, 1999, in: Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 749–783.
[12] J. Buzzi, V. Maume-Deschamps, Decay of correlations for piecewise invertible maps in higher dimensions, Isr. J. Math. 131 (2002) 203–220.
[13] S.G. Dani, On orbits of endomorphisms of tori and the Schmidt game, Ergod. Theory Dyn. Syst. 8 (4) (1988) 523–529.
[14] K.J. Falconer, Fractal Geometry - Mathematical Foundations and Applications, J. Wiley, Chichester, 1990.
[15] A.-H. Fan, Decay of correlation for expanding toral endomorphisms, in: Dynamical Systems, Proceedings of the International Conference in Honor of Professor Liao Shantao, World Scientific, 1999.
[16] A.-H. Fan, D.-J. Feng, J. Wu, Recurrence, dimension and entropy, J. Lond. Math. Soc. (2) 64 (1) (2001) 229–244.
[17] D. Färm, T. Persson, J. Schmeling, Dimension of countable intersections of some sets arising in expansions in non-integer bases, Fundam. Math. 209 (2) (2010) 157–176.
[18] H. Federer, Geometric Measure Theory, Springer-Verlag, 1969.
[19] J.L. Fernández, M.V. Melián, D. Pestana, Quantitative mixing results and inner functions, Math. Ann. 337 (2007) 233–251.
[20] P. Gallagher, Metric simultaneous Diophantine approximation, J. Lond. Math. Soc. (1962) 387–390.
[21] A.O. Gel'fond, A common property of number systems, Izv. Akad. Nauk SSSR, Ser. Mat. 23 (1959) 809–814.
[22] P. Góra, A. Boyarsky, Absolutely continuous invariant measures for piecewise expanding C^2 transformation in $\mathbb{R}^N$, Isr. J. Math. 67 (1989) 272–286.
[23] G. Harman, Metric Number Theory, LMS Monographs New Series, vol. 18, Clarendon Press, 1998.
[24] R. Hill, S.L. Velani, The ergodic theory of shrinking targets, Invent. Math. 119 (1995) 175–198.
[25] R. Hill, S.L. Velani, The shrinking target problem for matrix transformations of tori, Proc. Lond. Math. Soc. 60 (1999) 381–398.
[26] S. Ito, T. Sadahiro, Beta-expansions with negative bases, Integers 9 (A22) (2009) 239–259.
[27] C. Kalle, D.-R. Kong, N. Langeveld, W.-X. Li, The β-transformation with a hole at 0, Ergod. Theory Dyn. Syst. 40 (9) (2020) 2482–2514.
[28] G. Keller, Piecewise monotonic transformations and exactness, in: Seminar on Probability (Rennes, 1978), Exp. No. 6, Université de Rennes, Rennes, 1978, p. 32.
[29] G. Keller, Ergodicité et mesures invariantes pour les transformations dilatantes par morceaux d'une région bornée du plan, C. R. Acad. Sci. Paris, Ser. A-B 289 (1979) A625–A627.
[30] G. Keller, Propriété ergodique des endomorphismes dilatants, C^2 par morceaux, des régions bornées du plan, Thesis, Université de Rennes, 1979.
[31] B. Li, L. Liao, S. Velani, B.-W. Wang, E. Zorin, Tentative title: Simultaneous weighted Diophantine approximation with "rates" of infinite lower order, in preparation.

[32] B. Li, L. Liao, S. Velani, E. Zorin, The shrinking target problem for matrix transformations of tori: developing a manifold theory, in preparation.
[33] L. Liao, W. Steiner, Dynamical properties of the negative beta-transformation, Ergod. Theory Dyn. Syst. 32 (5) (2012) 1673–1690.
[34] D. Mauldin, M. Urbanski, Graph Directed Markov System: Geometry and Dynamics of Limit Sets, Cambridge Tracts in Mathematics, vol. 148, Cambridge University Press, 2003.
[35] P. Mattila, Geometry of sets and measures in Euclidean space, Camb. Stud. Adv. Math. 44 (1995).
[36] J. Nilsson, On numbers badly approximable by dyadic rationals, Isr. J. Math. 171 (2009) 93–110.
[37] M. Pollicott, M. Yuri, Dynamical Systems and Ergodic Theory, London Mathematical Society Student Texts, vol. 40, Cambridge University Press, Cambridge, 1998.
[38] F. Przytycki, M. Urbański, Conformal Fractals: Ergodic Theory Methods, London Mathematical Society Lecture Note Series, vol. 371, Cambridge University Press, Cambridge, 2010.
[39] A. Rényi, Representations for real numbers and their ergodic properties, Acta Math. Acad. Sci. Hung. 8 (1957) 477–493.
[40] V.A. Rokhlin, Exact endomorphisms of a Lebesgue space, Izv. Akad. Nauk SSSR, Ser. Mat. 25 (1961) 499–530.
[41] B. Rynne, Hausdorff dimension and generalized Diophantine approximation, Bull. Lond. Math. Soc. 30 (1998) 365–376.
[42] W. Philipp, Some metrical theorems in number theory, Pac. J. Math. 20 (1967) 109–127.
[43] B. Saussol, Absolutely continuous invariant measures for multidimensional expanding maps, Isr. J. Math. 116 (2000) 223–248.
[44] L.-M. Shen, B.-W. Wang, Shrinking target problems for beta-dynamical system, Sci. China Math. 56 (1) (2013) 91–104.
[45] V.G. Sprindžuk, Metric Theory of Diophantine Approximations, John Wiley, 1979, Translated by R. A. Silverman.
[46] S. Troubetzkoy, P. Varandas, The role of continuity and expansiveness on leo and periodic specification properties, 2020, hal-02557857v2.
[47] M. Tsujii, Absolutely continuous invariant measures for piecewise real-analytic expanding maps on the plane, Commun. Math. Phys. 208 (2000) 605–622.
[48] M. Tsujii, Absolutely continuous invariant measures for expanding piecewise linear maps, Invent. Math. 143 (2001) 349–373.
[49] M. Urbański, On Hausdorff dimension of invariant sets for expanding maps of a circle, Ergod. Theory Dyn. Syst. 6 (2) (1986) 295–309.
[50] G. Wagner, The ergodic behaviour of piecewise monotonic transformations, Z. Wahrscheinlichkeitstheor. Verw. Geb. 46 (3) (1979) 317–324.
[51] P. Walters, An Introduction to Ergodic Theory, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York-Berlin, 1982.
[52] B.-W. Wang, J. Wu, Mass transference principle from rectangles to rectangles in Diophantine approximation, Math. Ann. 381 (1–2) (2021) 243–317.
[53] Q.-Q. Yang, S.-L. Wang, Metric and dimensional properties of the badly approximable set for beta-transformations, Fractals 27 (3) (2019) 1950025, 9pp.